ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

THE FUNDAMENTAL THEOREM OF THE GALOIS THEORY FOR QUASI-FIELDS[1]

BY N. JACOBSON

(Received February 13, 1939)

It has been noted recently by Artin and by Baer that one obtains the fundamental theorem of the Galois theory most readily by starting with a finite group $\mathfrak{G}$ of automorphisms in a field P and determining the structure of P over Φ the set of invariant elements. One proves that P is finite, separable and normal and $\mathfrak{G}$ is its Galois group over Φ. The theorem is completed by proving that any isomorphism between subfields over Φ in a finite, separable and normal extension P of Φ can be extended to an automorphism in P. It follows that the Galois group of P over Φ has Φ as its set of invariant elements. The correspondence between subfields and subgroups follows readily. The usual proofs of these theorems are obtained by strictly commutative methods (symmetric functions, unique factorization of polynomials).

In the present paper we begin with an arbitrary quasi-field P and a finite group of outer automorphisms $\mathfrak{G}$ acting in P and establish the correspondence between subgroups of $\mathfrak{G}$ and sub-quasi-fields Σ between P and Φ the set of invariant elements. The methods are necessarily those of non-commutative algebra. The particular tool used is the theory of simple rings. We obtain some applications to division algebras. The first section is introductory containing results that are for the most part well known. In the last section we give a generalization of Hilbert's theorem on the elements of norm 1 in a cyclic field.

1. By a transformation T we shall understand a single-valued correspondence between the elements of a set P and those of a subset. We denote its effect on α by αT. T and S are regarded as equal only when they have the same effect on all α's of the set. The product ST is defined by $\alpha(ST) = (\alpha S)T$. If P is a group written additively $S + T$ is the transformation such that $\alpha(S + T) = \alpha S + \alpha T$ and T is an endomorphism if $(\alpha + \beta)T = \alpha T + \beta T$. With these conventions it is readily verified that the set of endomorphisms of a commutative group form a ring with an identity, namely, the identity transformation. The zero element is O given by $\alpha O = 0$ and the negative of T is $\alpha(-T) = -\alpha T$. If T is $(1 - 1)$ its inverse is also an endomorphism. When we speak of an endomorphism of a ring P we shall mean an endomorphism of its additive group.

A commutative group P in which there is defined a set of endomorphisms

[1] Presented to the Society April 7, 1939.

forming a quasi-field Φ which includes the identity transformation is called a vector space. P has finite dimensionality $(P:\Phi) = n$ if it contains n elements $\alpha_1, \alpha_2, \cdots, \alpha_n$ such that every element is expressible uniquely in the form $\alpha_1\xi_1 + \alpha_2\xi_2 + \cdots + \alpha_n\xi_n$, $\xi_i \in \Phi$. As is well known n is an invariant and may be characterized as the maximum number of linearly independent elements in P. We suppose hereafter that $(P:\Phi)$ is finite.

An endomoprhism A commutative with all the transformations in Φ is called a linear transformation (l. t.) of P over Φ. Their totality is a ring $\mathbf{A}$. Now suppose $\alpha_1, \alpha_2, \cdots, \alpha_n$ is a basis for P over Φ. Clearly A is determined by $\alpha_r A$ and we may choose $\alpha_r A$ arbitrarily and define a l. t. by the condition $(\sum \alpha_i\xi_i)A = \sum(\alpha_i A)\xi_i$. In particular we define E_{ij} by $\alpha_r E_{ij} = \delta_{ir}\alpha_j$ and ξ' by $\alpha_r\xi' = \alpha_r\xi$ for any ξ in Φ.[1a] The totality of transformations ξ' forms a quasi-field Φ' anti-isomorphic to Φ since $\alpha_r(\xi\eta)' = \alpha_r\xi\eta = \alpha_r\xi'\eta = \alpha_r\eta\xi' = \alpha_r\eta'\xi'$. Evidently P is a vector space over Φ' and since $\alpha_r\xi' = \alpha_r\xi, \alpha_1, \alpha_2, \cdots, \alpha_n$ form a basis for this space. It is readily seen that $\alpha_r\xi' E_{ij} = \alpha_r E_{ij}\xi'$ and hence $\xi' E_{ij} = E_{ij}\xi'$. If $\alpha_r A = \sum \alpha_j\xi_{jr}$ then $A = \sum E_{ij}\xi'_{ji}$ and hence the set of l. t.'s is identical with the set of these transformations. If $\sum E_{ij}\xi'_{ji} = 0$ then $\sum \alpha_j\xi_{jr} = 0$ and by the uniqueness we have $\xi_{jr} = 0$. If we regard P as a space over Φ' and use the basis α_i, we obtain $\alpha_r\xi'' = \alpha_r\xi$ and hence the l. t.'s ξ and ξ'' are identical. The E_{ij} are l. t.'s in P over Φ' and the general form of such a transformation is $\Sigma E_{ij}\xi_{ji}$. We note also that

$$E_{11} + E_{22} + \cdots + E_{nn} = 1, \qquad E_{ij}E_{kl} = \delta_{jk}E_{il}.$$

Hence if $E_{pq}(\Sigma E_{ij}\xi'_{ji}) = (\Sigma E_{ij}\xi'_{ji})E_{pq}$,

$$\Sigma E_{pj}\xi'_{jq} = \Sigma E_{iq}\xi'_{pi}$$

and $\xi'_{ij} = O$ if $i \neq j$, $\xi'_{ii} = \xi'$ and similarly the condition that $\Sigma E_{ij}\xi_{ji}$ commute with all E_{pq} is that it belong to Φ. Now if $T(T')$ is any endomorphism commutative with all the l. t.'s $\Sigma E_{ij}\xi'_{ji}(\Sigma E_{ij}\xi_{ji})$ it is an l. t. in P over $\Phi'(\Phi)$ and commutes with the E_{pq}. Hence $T \in \Phi(T' \in \Phi')$.

If P is a ring with an identity the transformations $\alpha \to \alpha\xi = \alpha\xi$ are endomorphisms and their totality is a ring $\mathbf{P}$ isomorphic to P. The transformations $\alpha \to \xi\alpha = \alpha\xi'$ are endomorphisms and form a ring $\mathbf{P}'$ anti-isomorphic to P. The elements of $\mathbf{P}$ commute with those of $\mathbf{P}'$. Furthermore if T' is any endomorphism commutative with all the ξ and $1T' = \tau$, then $\alpha T' = (1\alpha)T' = (1\alpha)T' = (1T')\alpha = \tau\alpha = \tau\alpha = \alpha\tau'$. Hence $T' \in P'$. Similarly $\mathbf{P}$ is the complete set of endomorphisms commutative with those of $\mathbf{P}'$.

If $P \geqq \Phi > 1$ and Φ is a quasi-field the transformations corresponding to the elements of Φ form a quasi-field Φ and P is a vector space over Φ. If left multiplication is used in place of right multiplication we obtain a quasi-field Φ' and P is a vector space over Φ'. Thus, for example, the set of l. t.'s of any vector space over Φ contains Φ' and may be regarded as a space over Φ'. Now suppose

[1a] Note that the ξ' as well as the E_{ij} depend on the basis $\alpha_1, \alpha_2, \cdots, \alpha_n$.

E is a space over $\mathbf{P}$ and $\mathbf{P} \geqq \Phi > 1$ where the dimensionalities $(\mathbf{E}:\mathbf{P})$ and $(\mathbf{P}:\Phi)$ are finite. If $\alpha_1, \alpha_2, \cdots, \alpha_n$ is a basis for E over $\mathbf{P}$, $\xi_1, \xi_2, \cdots, \xi_r$ one for $\mathbf{P}$ over Φ then $\beta_{ij} = \alpha_i\xi_j$ form a basis for E over Φ and hence $(\mathbf{E}:\Phi) = (\mathbf{E}:\mathbf{P})(\mathbf{P}:\Phi)$.

We recall also the definition of a semi-linear transformation $\boldsymbol{T}$ with automorphism $\boldsymbol{S}$ (in $\mathbf{P}$) as an endomorphism in E such that $(\alpha\xi)\boldsymbol{T} = (\alpha\boldsymbol{T})\xi^S$ for every ξ in $\mathbf{P}$.

2. Let P be an arbitrary quasi-field and $\mathfrak{G}$ a group of n automorphisms $\mathbf{1}$, $\boldsymbol{S}, \boldsymbol{T}, \cdots, \boldsymbol{U}$ in P $((\alpha + \beta)^S = \alpha^S + \beta^S, (\alpha\beta)^S = \alpha^S\beta^S)$. We suppose that no $\boldsymbol{S} \neq \mathbf{1}$ is inner, i.e. has the form $\alpha \to \xi^{-1}\alpha\xi$. Let $(\mathbf{P}, \boldsymbol{S})$ denote the set of endomorphisms of P generated by the transformations ξ of P, $\alpha \to \alpha\xi = \alpha\xi$ and the automorphisms $\boldsymbol{S}$. $(\mathbf{P}, \boldsymbol{S})$ consists of the sums $\Sigma\, \boldsymbol{S}\xi_S$ and $\xi\boldsymbol{S} = \boldsymbol{S}\xi^S$ where ξ^S corresponds to ξ^S. Suppose

$$\boldsymbol{S}\xi_S + \boldsymbol{T}\xi_T + \cdots + \boldsymbol{U}\xi_U = \boldsymbol{O}. \tag{1}$$

If $\xi_S \neq \boldsymbol{O} \neq \xi_T$ we multiply (1) on the left by $\mathfrak{n}$ and on the right by $\xi_S^{-1}\mathfrak{n}^S\xi_S$ and subtract to obtain

$$\boldsymbol{T}(\mathfrak{n}^T\xi_T - \xi_T\xi_S^{-1}\mathfrak{n}^S\xi_S) + \cdots = \boldsymbol{O}$$

which has at least one term less than (1). Since $\boldsymbol{T}\boldsymbol{S}^{-1}$ is not inner there exists an η such that

$$\eta^{TS^{-1}} \neq (\xi_S^{S^{-1}}(\xi_T^{S^{-1}})^{-1})^{-1}\eta(\xi_S^{S^{-1}}(\xi_T^{S^{-1}})^{-1})$$

and hence we may suppose that the new coefficient of $\boldsymbol{T}$ is $\neq 0$. Continuing in this way we get finally a single terms $\boldsymbol{S}\zeta_S = 0$ where $\zeta_S \neq 0$. Since $\boldsymbol{S}$ and ζ_S have inverses this is impossible. Thus the space $(\mathbf{P}, \boldsymbol{S})$ over $\mathbf{P}$ has n dimensions.

$(\mathbf{P}, \boldsymbol{S})$ is a simple ring with an identity. For if $\mathfrak{B}$ is a two-sided ideal and contains $\sum \boldsymbol{S}\xi_S \neq \boldsymbol{O}$ the above argument shows that $\mathfrak{B}$ contains some $\boldsymbol{S}\zeta_S$, $\zeta_S \neq \boldsymbol{O}$. Then $\mathfrak{B} > \mathbf{1}$ and $\mathfrak{B} = (\mathbf{P}, \boldsymbol{S})$. Any right ideal $\mathfrak{J}$ of $(\mathbf{P}, \boldsymbol{S})$ is a subspace over $\mathbf{P}$. It follows that $(\mathbf{P}, \boldsymbol{S})$ satisfies the chain conditions for right ideals. Hence the following theorems hold.[2]

Any two irreducible right ideals are similar, i.e. there exists a (1–1) correspondence $\boldsymbol{X} \to \overline{\boldsymbol{X}}$ between $\mathfrak{J}$ and $\overline{\mathfrak{J}}$ such that

$$\overline{\boldsymbol{X} + \boldsymbol{Y}} = \overline{\boldsymbol{X}} + \overline{\boldsymbol{Y}} \qquad \overline{\boldsymbol{X}\boldsymbol{A}} = \overline{\boldsymbol{X}}\boldsymbol{A}$$

for any $\boldsymbol{A}$ in $(\mathbf{P}, \boldsymbol{S})$. In particular, if we restrict $\boldsymbol{A}$ to be in $\mathbf{P}$ we see that $\boldsymbol{X} \to \overline{\boldsymbol{X}}$ is a (1–1) l. t. between $\mathfrak{J}$ and $\overline{\mathfrak{J}}$ and hence $(\mathfrak{J}:\mathbf{P}) = (\overline{\mathfrak{J}}:\mathbf{P})$. Since the elements $(\mathbf{1} + \boldsymbol{S} + \boldsymbol{T} + \cdots + \boldsymbol{U})\xi$, ξ variable, form a right ideal of dimension 1, we see that $(\mathfrak{J}:\mathbf{P}) = 1$ for every irreducible $\mathfrak{J}$.

$(\mathbf{P}, \boldsymbol{S}) = \mathfrak{J}_1 + \mathfrak{J}_2 + \cdots + \mathfrak{J}_r$ (direct sum) where the $\mathfrak{J}_i$ are irreducible right ideals. Hence $n = ((\mathbf{P}, \boldsymbol{S}):\mathbf{P}) = r(\mathfrak{J}_i:\mathbf{P}) = r$. It follows that $(\mathbf{P}, \boldsymbol{S}) = \Psi_n$ a

[2] See v. d. Waerden, *Moderne Algebra*, II, pp. 149–177 or Deuring, *Algebren* pp. 8–25.

matrix ring, i.e. the elements of $(\mathbf{P}, \boldsymbol{S})$ have the form $\Sigma \boldsymbol{E}_{ij}\psi_{ij}$ where the ψ's form a quasi-field, $\boldsymbol{E}_{ij} \neq \mathbf{0}$, commute with the ψ's and satisfy

$$\boldsymbol{E}_{11} + \boldsymbol{E}_{22} + \cdots + \boldsymbol{E}_{nn} = \mathbf{1} \qquad \boldsymbol{E}_{ij}\boldsymbol{E}_{kl} = \delta_{jk}\boldsymbol{E}_{il}\,. \tag{2}$$

Let α be any element $\neq 0$ in P. By (2) there is an $\boldsymbol{E}_{pp}$ such that $\alpha\boldsymbol{E}_{pp} \neq 0$ and since $\alpha\boldsymbol{E}_{pp} = \alpha\boldsymbol{E}_{pi}\boldsymbol{E}_{ip}$, this implies that no $\alpha_i = \alpha\boldsymbol{E}_{pi}$ is 0. Suppose

$$\alpha_1\psi_1 + \alpha_2\psi_2 + \cdots + \alpha_n\psi_n = 0.$$

Then $(\alpha\boldsymbol{E}_{p1}\psi_1 + \alpha\boldsymbol{E}_{p2}\psi_2 + \cdots + \alpha\boldsymbol{E}_{pn}\psi_n)\boldsymbol{E}_{ii} = (\alpha\boldsymbol{E}_{pi})\psi_i = 0$ which implies $\psi_i = 0$. On the other hand since $\boldsymbol{\Psi}_n$ includes $\mathbf{P}$, any element γ in P can be obtained from a fixed $\beta \neq 0$ by operating with a suitable $\sum \boldsymbol{E}_{ij}\psi_{ij}$. Thus

$$\gamma = (\alpha\boldsymbol{E}_{pp})\sum \boldsymbol{E}_{ij}\psi_{ij} = \alpha_1\psi_1 + \alpha_2\psi_2 + \cdots + \alpha_n\psi_n\,.$$

Hence the α_i form a basis for P over $\boldsymbol{\Psi}$ and from their definition we have $\alpha_r\boldsymbol{E}_{ij} = \delta_{ir}\alpha_j$.

As in §1 we associate with each ψ the transformation ψ' defined by $(\sum \alpha_i\psi_i)\psi' = \sum \alpha_i\psi\psi_i$ and denote the totality of these transformations as $\boldsymbol{\Psi}'$. P is a vector space over $\boldsymbol{\Psi}'$ with a basis $\alpha_1, \alpha_2, \cdots, \alpha_n$ as a basis. The transformations $\sum \boldsymbol{E}_{ij}\psi_{ij}$ form the complete set of l. t.'s in P over $\boldsymbol{\Psi}'$ and the elements of $\boldsymbol{\Psi}'$ may be characterized as the endomorphisms commutative with all the $\sum \boldsymbol{E}_{ij}\psi_{ij}$.

On the other hand $\boldsymbol{\Psi}_n = (\mathbf{P}, \boldsymbol{S})$ and if an endomorphism commutes with all of $\mathbf{P}$ it is in $\mathbf{P}'$ the set of left multiplications $\alpha \to \xi\alpha = \alpha\xi'$ and if in addition it commutes with all the $\boldsymbol{S}$'s we must have ξ in Φ. We have therefore shown that $\boldsymbol{\Psi}' = \boldsymbol{\Phi}'$ and $(\mathbf{P}, \boldsymbol{S})$ is the complete set of l. t.'s of P over $\boldsymbol{\Phi}'$. If we use $(\mathbf{P}', \boldsymbol{S})$ in place of $(\mathbf{P}, \boldsymbol{S})$ we obtain $(\mathrm{P}:\boldsymbol{\Phi}) = n$ and $(\mathbf{P}', \boldsymbol{S})$ is the complete set of l. t.'s in the vector space P over $\boldsymbol{\Phi}$. Hence we have the important theorem.

THEOREM 1. *Let* P *be an arbitrary quasi-field,* $\mathfrak{G}$ *a finite group of* n *outer automorphisms acting in* P *and* Φ *the sub-quasi-field of invariant elements. Then the order* $(\mathrm{P}:\boldsymbol{\Phi}) = (\mathrm{P}:\boldsymbol{\Phi}') = n$ *and* $(\mathbf{P}, \boldsymbol{S})$ *the set of transformations* $\Sigma \boldsymbol{S}\xi_{\boldsymbol{S}}$ *is the complete set of linear transformations of* P *over* $\boldsymbol{\Phi}'$, $(\mathbf{P}', \boldsymbol{S})$ *the complete set of linear transformations of* P *over* $\boldsymbol{\Phi}$.

Suppose $\boldsymbol{V}$ is any automorphism in P leaving the elements of Φ unaltered. Then $\boldsymbol{V}$ is a l. t. in P over $\boldsymbol{\Phi}'$ and hence $\boldsymbol{V} = \sum \boldsymbol{S}\xi_{\boldsymbol{S}}$. For every η we have $\boldsymbol{n}\boldsymbol{V} - \boldsymbol{V}\boldsymbol{n}^{\boldsymbol{V}} = \boldsymbol{O}$. Hence $\sum \boldsymbol{S}(\boldsymbol{n}^{\boldsymbol{S}}\xi_{\boldsymbol{S}} - \xi_{\boldsymbol{S}}\boldsymbol{n}^{\boldsymbol{V}}) = \boldsymbol{O}$ and if $\xi_{\boldsymbol{S}} \neq 0$, $\eta^{\boldsymbol{V}} = \xi_{\boldsymbol{S}}^{-1}\eta^{\boldsymbol{S}}\xi_{\boldsymbol{S}}$. Since no $\boldsymbol{S} \neq \mathbf{1}$ is inner this holds for just one $\boldsymbol{S}$ and $\boldsymbol{V} = \boldsymbol{S}\boldsymbol{A}$ where $\boldsymbol{A}$ is inner. $\boldsymbol{A}$ must leave the elements of Φ invariant. Hence if $\alpha^{\boldsymbol{A}} = \xi^{-1}\alpha\xi$, ξ commutes with all the α's in Φ.

If $\mathfrak{H}$ is a subgroup of $\mathfrak{G}$ we denote the sub-quasi-field of elements invariant under the transformations of $\mathfrak{H}$ by $\mathrm{P}(\mathfrak{H})$ and if Σ is any sub-quasi-field between Φ and P we denote the subgroup of $\mathfrak{G}$ leaving the elements of Σ invariant by $\mathfrak{G}(\Sigma)$. Note that $\Phi \leqq \mathrm{P}(\mathfrak{H}) \leqq \mathrm{P}$, $\mathrm{P}(\mathfrak{G}) = \Phi$ $\mathrm{P}(\mathbf{1}) = \mathrm{P}$. The following is the fundamental theorem of the Galois theory.

THEOREM 2. *The correspondences* $\mathfrak{H} \to \mathrm{P}(\mathfrak{H})$ *and* $\Sigma \to \mathfrak{G}(\Sigma)$ *are inverses of each other. Either one sets up a* (1–1) *correspondence between the subgroups of* $\mathfrak{G}$

and the quasi-fields Σ *between* Φ *and* P. *The order* $(P:\Sigma) = (P:\Sigma') =$ *order of* $\mathfrak{G}(\Sigma)$ *and* $(\Sigma:\Phi) = (\Sigma:\Phi') =$ *index of* $\mathfrak{G}(\Sigma)$.

Let $\mathfrak{H}$ be a subgroup of $\mathfrak{G}$ and $P(\mathfrak{H})$ the set of invariant elements. If T is an automorphism of $\mathfrak{G}$ leaving the elements of $P(\mathfrak{H})$ invariant, $T = SA$ where S is in $\mathfrak{H}$ and A is inner. Since $A = S^{-1}T$ is in $\mathfrak{G}$ this implies that $A = 1$, $S = T \in \mathfrak{H}$. Thus $\mathfrak{G}(P(\mathfrak{H})) = \mathfrak{H}$. Now suppose Σ is given where $\Phi \leqq \Sigma \leqq P$ and let Λ be the set of l. t.'s of P over Σ'. Then $\Lambda \leqq (P, S)$. If $\Sigma S \xi_S \in \Lambda$

$$\sum S\mu'^S\xi_S = \mu' \sum S\xi_S = (\sum S\xi_S)\mu' = \sum S\mu'\xi_S$$

Hence $\sum S(\mu'^S - \mu')\xi_S = O$. The process used above shows that $\xi_S = O$ unless $\mu^S = \mu$ for all μ in Σ, i. e. unless $S \in \mathfrak{G}(\Sigma) = \mathfrak{H}$. Hence Λ consists of the transformations $\sum_{S \in \mathfrak{H}} S\xi_S$. Σ' is the complete set of endomorphisms commutative with those of Λ. On the other hand the form of the elements of Λ shows that these transformations are precisely the ζ's such that $\zeta \in P(\mathfrak{H})$. Hence $P(\mathfrak{G}(\Sigma)) = \Sigma$. The order relations follow from Theorem 1.

If $\Sigma = P(\mathfrak{H})$, $\Sigma^S = P(S^{-1}\mathfrak{H}S)$. Hence $\mathfrak{H}$ is invariant if and only if Σ is transformed into itself by all the elements of $\mathfrak{G}$. If $\bar{1}, \bar{S}, \cdots$ are representatives of the cosets of $\mathfrak{H}$, the transformations induced in Σ by these elements are distinct and depend only on the cosets. Their totality is a group $\bar{\mathfrak{G}} \cong \mathfrak{G}/\mathfrak{H}$.

3. If P is commutative, ξ any element in P, let $\xi, \xi^S, \cdots, \xi^T$ be its distinct conjugates. The coefficients of

$$(x - \xi)(x - \xi^S) \cdots (x - \xi^T)$$

are invariant under $\mathfrak{G}$ and therefore belong to Φ. Hence every element of P satisfies an equation with coefficients in Φ and with distinct roots in P. It follows that P is a separable, normal extension of Φ. To complete the Galois theory for finite extensions along these lines one needs only the converse theorem that if P is a finite, separable, normal extension of Φ, then the Galois group leaves invariant only the elements of Φ.[3]

Let Γ be the center of P and $\Delta = \Gamma \cap \Phi$. The automorphisms S induce automorphisms S' in Γ and Δ is the set of invariant elements relative to $\mathfrak{G}'$ the set of S''s. Suppose first that the order of $\mathfrak{G}'$ = order of $\mathfrak{G} = n$. $\Phi\Gamma$ the join of Φ and Γ is invariant relative to $\mathfrak{G}$ only. Hence $\Phi\Gamma = P$. If $\gamma_1, \gamma_2, \cdots, \gamma_n$ is a basis for Γ over Δ every element of P has the form $\gamma_1\varphi_1 + \gamma_2\varphi_2 + \cdots + \gamma_n\varphi_n$, φ_i in Φ and since $(P:\Phi) = n$, these elements form a basis for P over Φ. Hence if we regard P, Φ and Γ as algebras over Δ we have $P = \Phi \times \Gamma$.[4] It is then evident that the center of Φ is Δ, i. e. Φ is a normal division algebra over Δ.[5]

---

[3] v. d. Waerden, *Moderne Algebra* I, p. 150, 1 st ed. or p. 161, 2nd ed.

[4] If P is an algebra over Δ with an identity 1, Φ and Γ subalgebras containing 1, $P = \Phi \times \Gamma$ means that 1) $P = \Phi\Gamma$, 2) the elements of Φ commute with those of Γ and 3) if $\varphi_1, \varphi_2, \cdots, \varphi_m$ are linearly independent in Φ over Δ and $\gamma_1, \gamma_2, \cdots, \gamma_r$ are independent in Γ over Δ then $\varphi_i\gamma_j$ are independent in P over Δ.

[5] Since $P = \Phi \times \Gamma$ any element commutative with all $\varphi \in \Phi$ is in Γ. It follows that if V is any automorphism leaving the elements of Φ invariant then $V \in \mathfrak{G}$.

The converse of this is clear. If $P = \Phi \times \Gamma$ where Φ is a normal division algebra over Δ and Γ is a separable, normal field over Δ then the Galois group of Γ has order $n = (\Gamma:\Delta)$. These automorphisms have a natural extension in $P = \Phi \times \Gamma$ obtained by requiring that they leave the elements of Φ unaltered. The automorphisms thus obtained are outer.

THEOREM 3. *Let* P *be a quasi-field,* Γ *its center. A necessary and sufficient condition that* P *have a group of* n *outer automorphisms which induce distinct automorphisms in* Γ *is that* $P = \Phi \times \Gamma$ *where* Φ *is a normal division algebra over* Δ *a subfield of* Γ *and* Γ *is separable and normal of order* n *over* Δ.

If $(P:\Gamma)$ is finite any automorphism leaving the elements of Γ invariant is inner.[6] Hence in this case the group $\mathfrak{G}'$ is necessarily of order n.

THEOREM 4. *If* P *is a quasi-field such that* $(P:\Gamma)$ *is finite, a necessary and sufficient condition that* P *have a group of* n *outer automorphisms is that* $P = \Phi \times \Gamma$ *where* Φ *is a normal division algebra over* Δ *a subfield of* Γ *and* Γ *is separable and normal of order* n *over* Δ.

Now suppose Γ is a $\mathfrak{p}$-adic field[7] and $(P:\Gamma) = m^2$. Then $(\Phi:\Delta) = m^2$ and by a known theorem on $\mathfrak{p}$-adic algebras $\Phi \times \Gamma$ is a division algebra if and only if $(m, n) = 1$.[8] Hence we have the theorem.

THEOREM 5. *If* P *is a quasi-field with center* Γ *a* $\mathfrak{p}$-adic *field and* $(P:\Gamma) = m^2$, *then any group of outer automorphisms has order prime to* m.

In particular P has no outer automorphism S of order l a prime divisor of m since $S, S^2, \cdots, S^l$ would then be outer automorphisms.

We return now to the general case. Let $\Delta = \Gamma \cap \Phi$, $\mathfrak{G}'$ the group of induced automorphisms and $\mathfrak{H}$ the subgroup of elements T such that $\gamma^T = \gamma$ for all γ in Γ. $\mathfrak{H}$ is an invariant subgroup of $\mathfrak{G}$ and $\mathfrak{G}' \cong \mathfrak{G}/\mathfrak{H}$. If $\Sigma = P(\mathfrak{H})$, $P \geqq \Sigma \geqq \Phi$ and the automorphisms S of $\mathfrak{G}$ induce automorphisms $\bar{S}$ in Σ which form a group $\bar{\mathfrak{G}} \cong \mathfrak{G}/\mathfrak{H}$. Since $\Sigma \geqq \Gamma$ the $\bar{S}$ are outer and induce distinct automorphisms S' in Γ. Theorem 3 may then be applied to Σ. We shall not attempt a more refined analysis of the structure of P here. We merely give a very simple example to show that P may be $> \Sigma$.

4. The following example due to Köthe shows that the automorphisms of $\mathfrak{G}$ need not be distinct in Γ.[9] Let Γ_0 be any field of characteristic $\neq 2$ and $\Gamma = \Gamma_0(x_1, y_1, x_2, y_2, \cdots)$ where the x_i and y_i are distinct indeterminates. Let $P^{(m)} = (x_1, y_1) \times (x_2, y_2) \times \cdots \times (x_m, y_m)$ where (x_i, y_i) is the generalized quaternion algebra with basis $1, \xi_i, \eta_i, \xi_i\eta_i$ such that

$$\xi_i^2 = x_i, \qquad \eta_i^2 = y_i, \qquad \xi_i\eta_i = -\eta_i\xi_i.$$

$P^{(m)}$ is a quasi-field and hence so is $P = \sum_1^\infty P^{(m)}$. If $\zeta_m = \xi_1\xi_2 \cdots \xi_m$ then $\alpha_m \to \zeta_m^{-1}\alpha_m\zeta_m$ is an automorphism S_m of period 2 in $P^{(m)}$ leaving the elements

[6] Deuring p. 43.

[7] i.e. a finite extension of the field of ordinary p-adic numbers of a field of power series in one indeterminate with coefficients in a finite field.

[8] Deuring p. 113 and p. 137.

[9] "Schiefkörper unendlichen Ranges über den Zentrum," *Math. Annalen* 105, p. 24.

of Γ unaltered. S_m has the same effect as S_{m-1} in $P^{(m-1)}$. Hence $S = \lim S_m$ is defined and is an automorphism in P. S is not inner, for if $\alpha^S = \zeta^{-1}\alpha\zeta$, ζ is in $P^{(N)}$ for N sufficiently large and hence

$$\eta_{N+1}{}^S = (\xi_1 \cdots \xi_{N+1})^{-1}\eta_{N+1}(\xi_1 \cdots \xi_{N+1})^{-1} = -\eta_{N+1} = \zeta^{-1}\eta_{N+1}\zeta = \eta_{N+1}$$

which is impossible. $S^2 = 1$ and $S = 1$ in Γ. Since $\zeta_m^2 = x_1x_2 \cdots x_m$ the dimensionality over Γ of $\Phi^{(m)} = P^{(m)} \cap \Phi$ is 2^{2m-1}.[10] It follows that 1, η_1 is a basis for $P^{(m)}$ over $\Phi^{(m)}$ and hence for P over Φ.

5. Let E be a vector space over $\mathbf{P}$, $(E:P) = r$ and $T_1 = 1, T_S, \cdots, T_U$ a finite group of n semi-linear transformations such that the automorphisms S in $\mathbf{P}$ associated with T_S ($\xi T_S = T_S\xi^S$) are distinct and outer. Let Φ denote the set of invariant elements of $\mathbf{P}$ and H those of E. H is a space over Φ. The transformations $\sum T_S\xi_S$ form a ring isomorphic to $(\mathbf{P}, S)$ defined above.[11] Hence these transformations may also be represented in the form $\sum E_{ij}\psi_{ij}$ as in §2. If $\zeta = \epsilon E_{pp} \neq 0$ then any vector in $E' = \zeta \sum T_S\xi_S$ has the form $\zeta_1\psi_1 + \zeta_2\psi_2 + \cdots + \zeta_n\psi_n$ where $\zeta_i = \zeta E_{pi}$ and $\psi_i \in \Psi$. It follows that if $\mu \neq 0$ is any vector in E', $\mu\Psi_n = E'$. The transformations induced by the $\sum T_S\xi_S$ in E' form a ring isomorphic to $(\mathbf{P}, S)$ also. Since $E = \sum T_S \neq 0$ there is a μ in E' such that $\mu E = \eta \neq 0$. Then $\eta \in$ H and $\zeta = \eta \sum T_S\xi_S = \eta\xi$, ξ in $\mathbf{P}$. Since any $\epsilon = \epsilon E_{11} + \epsilon E_{22} + \cdots + \epsilon E_{nn}$, ϵ is a linear combination of elements η_i in H with coefficients in $\mathbf{P}$. Hence E has a basis $\eta_1, \eta_2, \cdots, \eta_r$ over $\mathbf{P}$. If $\eta = \Sigma\eta_i\xi_i$ is in H it follows that $\xi_i \in \Phi$ and $\eta_1, \eta_2, \cdots, \eta_r$ form a basis for H over Φ.

THEOREM 6. *Let* E *be a vector space over* $\mathbf{P}$, $(E:P) = r$ *and* $T_1 = 1, T_S, \cdots, T_U$ *a finite group of semi-linear transformations whose induced automorphisms* $1, S, \cdots, U$ *in* $\mathbf{P}$ *are distinct and outer. If* H *is the set of vectors invariant under the* T_S, Φ *the set of elements invariant under the* S, *then* H *is a space over* Φ *such that* $(H:\Phi) = r$ *and the extension* $HP = E$.

We recall that if T_S is the matrix of T_S relative to the basis $\epsilon_1, \epsilon_2, \cdots, \epsilon_r$, i. e. $(\epsilon_1T_S, \epsilon_2T_S, \cdots, \epsilon_rT_S) = (\epsilon_1, \epsilon_2, \cdots, \epsilon_r)T_S$, then the matrix relative to a second basis $(\epsilon_1, \epsilon_2, \cdots, \epsilon_r)A$ is $A^{-1}T_SA^S$.[12] The conditions $T_ST_T = T_U$ when $ST = U$ are equivalent to $T_TT_S^T = T_{ST}$. Since the matrices of the T_S are all 1 relative to $(\eta_1, \eta_2, \cdots, \eta_r)$ we have proved

THEOREM 7. *If* $\mathfrak{G}$ *is a finite group of outer automorphisms in a quasi-field* P *and* T_S *are matrices with elements in* P *such that* $T_1 = 1$, $T_ST_S^T = T_{ST}$, *then there exists a non-singular matrix* A *such that* $T_S = A^{-1}A^S$ *for all* S.

UNIVERSITY OF NORTH CAROLINA.

[10] If K is a subfield of a normal division algebra P and Λ is the set of elements commutative with those of K, then $(P:\Gamma) = (K:\Gamma)(\Lambda:\Gamma)$. See Deuring p. 44.

[11] Evidently $\sum S\xi_S \to \sum T_S\xi_S$ is a homomorphism. Since $(\mathbf{P}, S)$ is simple this correspondence is an isomorphism.

[12] Jacobson, "Pseudo-linear transformations," *Annals of Math.* 38 (1937) p. 486.

ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

ON A GEOMETRICAL REPRESENTATION OF p-ADIC NUMBERS

BY KURT MAHLER

(Received December 30, 1938)

Let ζ be a real irrational number,

$$\zeta = a_0 + \frac{1|}{|a_1} + \frac{1|}{|a_2} + \cdots$$

its simple continued fraction, and

$$\frac{p_0}{q_0} = \frac{a_0}{1}, \qquad \frac{p_1}{q_1} = \frac{a_0 a_1 + 1}{a_1}, \qquad \frac{p_2}{q_2}, \cdots$$

the sequence of its approximations. The following theorems have been proved:

THEOREM OF LAGRANGE:[1] *If p/q is a rational number, such that*

$$\left|\zeta - \frac{p}{q}\right| \leqq \left|\zeta - \frac{p_n}{q_n}\right|$$

for a certain index $n \geqq 1$, then $|q| \geqq q_n$, with equality if and only if $\frac{p}{q} = \frac{p_n}{q_n}$.

THEOREM OF HURWITZ-BOREL:[2] *For at least one of any three consecutive indices n*

$$\left|\zeta - \frac{p_n}{q_n}\right| < \frac{1}{\sqrt{5}}\, q_n^{-2},$$

but to every $\epsilon > 0$, there is an irrational number ζ, such that

$$\left|\zeta - \frac{p_n}{q_n}\right| \geqq \left(\frac{1}{\sqrt{5}} - \epsilon\right) q_n^{-2}$$

for all sufficiently large n.

THEOREM OF KHINTCHINE:[3] *There are arbitrarily large positive integers t, for which the inequalities*

$$\left|\zeta - \frac{p}{q}\right| < \frac{1}{t\,|q|}, \qquad 0 < |q| \leqq \frac{t}{2}$$

have no solution in integers p, q.

---

[1] O. Perron, *Die Lehre von den Kettenbrüchen*, Leipzig-Berlin 1929, §15.
[2] l. c. 1, §14.
[3] J. F. Koksma, *Diophantische Approximationen*, Berlin 1936, p. 36 f.

THEOREM OF TCHEBYCHEFF:[4] *If ζ is an irrational, ϑ an arbitrary real number, then there are arbitrarily large positive integers t, for which the inequalities*

$$|q\zeta - p - \vartheta| \leqq \frac{1}{t}, \qquad |q| \leqq \frac{t}{2}$$

have an integer solution p, q.

(Though the statement of the last two theorems does not mention continued fractions, their proof is much simpler if these are used.)

In the present paper, I derive analogous results for P-adic numbers, when P is an arbitrary positive prime number; see the Theorems 17–26 in Part II. For this purpose it is necessary to develop a P-adic algorithm similar to that of the continued fractions in the real field. In two earlier papers,[5] I have previously studied and applied an algorithm of this kind; the method there was purely arithmetical and based on Minkowski's Theorem on linear forms. It had, however, some disadvantages, e.g. it did not lead to analogues to the Theorems of Hurwitz-Borel and Khintchine with good values of the constants.

For this reason, I shall use in this paper a geometrical method, by which the best approximations of a given P-adic integer ζ can always be obtained. (The restriction to P-adic integers is unimportant, since, if necessary, $1/\zeta$ instead of ζ may be considered.) This method forms something like a P-adic counterpart to Hermite's method of introduction of continuous variables in the theory of forms.[6] It is based on the following idea:

Let

$$\mathfrak{T}_n = \begin{pmatrix} p_n & p_n' \\ q_n & q_n' \end{pmatrix} \qquad (n = 0, 1, 2, \ldots)$$

be an infinite sequence of integer matrices with the following properties:

(a): $$p_n q_n' - p_n' q_n = P^n.$$

(b): $$(q_n, q_n') = 1.$$

(c): All matrices $\mathfrak{T}_n^{-1}\mathfrak{T}_{n+1} = \Omega_{n+1}$ have integer elements.

Then it is easily verified that to every n there is an integer A_n, such that $0 \leqq A_n \leqq P^n - 1$, $p_n + q_n A_n \equiv 0 \pmod{P^n}$, $p_n' + q_n' A_n \equiv 0 \pmod{P^n}$, and such that

$$A_{n+1} \equiv A_n \pmod{P^n}.$$

There exists, therefore, a P-adic integer ζ as the P-adic limit

$$\zeta = \lim_{n \to \infty} A_n.$$

---

[4] l. c. 3, p. 76 f.

[5] *Nieuw Arch. Wiskde* (2), **18** (1934), 22–34, and *Mathematica* B (Zutphen), VII (1938), 5 p.

[6] Ch. Hermite, *Oeuvres*, vol. 1, Paris 1905, 100–163, 164–193, 200–263, and l. c. 3, p. 40 f.

This number is determined uniquely by the sequence $\{\mathbf{T}_n\}$, and it is related to its elements $\mathbf{T}_n$ by the inequalities

$$|p_n + q_n\zeta|_P \leqq P^{-n}, \qquad |p'_n + q'_n\zeta|_P \leqq P^{-n}$$

and the equations[7]

$$\zeta = \lim_{n\to\infty}\left(-\frac{p_n}{q_n}\right) \quad \text{or} \quad \lim_{n\to\infty}\left(-\frac{p'_n}{q'_n}\right).$$

We say that the sequence $\{\mathbf{T}_n\}$ defines ζ. Conversely, every P-adic integer can be defined by a sequence $\{\mathbf{T}_n\}$; it is sufficient to take

$$\mathbf{T}_n = \begin{pmatrix} P^n & -A_n \\ 0 & 1 \end{pmatrix} \qquad (n = 0, 1, 2, \cdots),$$

where A_n is any integer for which

$$|\zeta - A_n|_P \leqq P^{-n}.$$

Then $(A_{n+1} - A_n)/P^n$ is an integer, and therefore also all matrices

$$\mathbf{T}_n^{-1}\mathbf{T}_{n+1} = \begin{pmatrix} -P & (A_{n+1} - A_n)P^{-n} \\ 0 & -1 \end{pmatrix}$$

are integral. The three conditions (a), (b), (c) are therefore satisfied.

Let us consider two sequences $\{\mathbf{T}_n\}$ and $\{\mathbf{T}_n^*\}$ as equivalent if they define the same number ζ. This is the case if and only if to every index $n = 0, 1, 2, \cdots$ there is an element $\mathbf{P}_n$ of the modular group (i.e. an integer matrix of determinant 1), such that

$$\mathbf{T}_n^* = \mathbf{T}_n\mathbf{P}_n .$$

Among all equivalent sequences $\{\mathbf{T}_n\}$, which define ζ, we can choose one as the reduced sequence. This we may do, for instance, in the following way:

Let $\Phi(X, Y)$ be the distance function of an arbitrary convex domain $\Phi(X, Y) \leqq 1$ of area J in the (X, Y)-plane with center at the origin; thus

$$\Phi(0, 0) = 0, \qquad \Phi(X, Y) > 0 \quad \text{for} \quad X^2 + Y^2 > 0,$$

$$\Phi(tX, tY) = |t|\,\Phi(X, Y) \qquad \text{for real } t,$$

$$\Phi(X_1 + X_2, Y_1 + Y_2) \leqq \Phi(X_1, Y_1) + \Phi(X_2, Y_2).$$

With every element $\mathbf{T}_n = \begin{pmatrix} p_n & p'_n \\ q_n & q'_n \end{pmatrix}$ of $\{\mathbf{T}_n\}$ we form the new convex function

$$\Phi_n(X, Y) = \Phi(p_nX + p'_nY, q_nX + q'_nY);$$

[7] One of these limits may be different from ζ.

the domain $\Phi_n(X, Y) \leqq 1$ is again convex with center at the origin, and its area is $P^{-n}J$. Suppose now that for every $n = 0, 1, 2, \cdots$ the function $\Phi_n(X, Y)$ satisfies the inequalities

(d): $$\Phi_n(1, 0) \leqq \Phi_n(0, 1) \leqq \Phi_n(-1, 1) \leqq \Phi_n(1, 1);$$

then we say that the sequence $\{\mathbf{T}_n\}$ is reduced. By Minkowski,[8] it follows from (d) that

(e): $$\frac{2P^n}{J} \leqq \Phi_n(1, 0)\Phi_n(0, 1) \leqq \frac{4P^n}{J}.$$

Hence in particular, for all indices n simultaneously

(f): $$|p_n + q_n\zeta|_P \leqq P^{-n}, \qquad \Phi(p_n, q_n) \leqq \frac{2P^{n/2}}{\sqrt{J}}.$$

It is not difficult to show that there is one and only one reduced sequence which defines ζ; by (f), the elements of this sequence lead to approximations $-p_n/q_n$ for ζ.—It can also be shown that if $\{\mathbf{T}_n\}$ is reduced, then every term of the allied sequence

$$\Omega_{n+1} = \mathbf{T}_n^{-1}\mathbf{T}_{n+1} \qquad (n = 0, 1, 2, \cdots)$$

belongs to a finite set of matrices, which depends only on P and on the function $\Phi(X, Y)$, but not on ζ. By means of this result, it is possible to determine indices n, for which the inequalities (f) can be improved, provided that such indices exist; it is further possible to obtain results analogous to the Theorem of Khintchine.

This is carried through in the present paper for the case in which

$$\Phi(X, Y)^2 = \mathbf{A}X^2 + 2\mathbf{B}XY + \mathbf{C}Y^2$$

is a positive definite quadratic form of determinant 1. In this case, the inequalities (e) and (f) can be improved a little. It is more important, however, that the conditions (d) now become the well known conditions for a reduced positive definite quadratic form, viz.

$$-\mathbf{A}_n \leqq 2\mathbf{B}_n < \mathbf{A}_n \leqq \mathbf{C}_n,$$

if

$$\Phi_n(X, Y)^2 = \mathbf{A}_nX^2 + 2\mathbf{B}_nXY + \mathbf{C}_nY^2.$$

Therefore, if in the usual way[9] we represent the form $\Phi_n(X, Y)$ by the point z_n in the upper half of the complex plane which satisfies the equation $\Phi_n(z, 1) = 0$, then z_n will lie in the fundamental domain F of the modular group. Hence,

[8] *Geometrie der Zahlen*, Leipzig-Berlin 1910, 193–196.

[9] P. Bachmann, *Quadratische Formen II*, Leipzig-Berlin 1923, p. 17 f.

as a geometrical representative of ζ, we get an infinite set of points z_n in F. Two consecutive points of this set are connected by the equation

$$(g): \qquad z_n = \frac{\alpha_{n+1} z_{n+1} + \alpha'_{n+1}}{\beta_{n+1} z_{n+1} + \beta'_{n+1}},$$

if $\Omega_{n+1} = \begin{pmatrix} \alpha_{n+1} & \alpha'_{n+1} \\ \beta_{n+1} & \beta'_{n+1} \end{pmatrix}$. As we have already remarked, the matrices Ω_{n+1} have only a finite number of possibilities independent of ζ; it is quite easy to find all these for a given prime number P. We can, therefore, study the elements of the sequence $\{z_n\}$ without specializing ζ; for instance, we can find lower bounds for $\limsup_{n\to\infty} y_n$, where $z_n = x_n + iy_n$. Since

$$| p_n + q_n\zeta |_P \leqq P^{-n}, \qquad | p'_n + q'_n\zeta |_P \leqq P^{-n}$$

and

$$\Phi(p_n, q_n)^2 = \frac{P^n}{y_n}, \qquad \Phi(p'_n, q'_n) = \frac{P^n(x_n^2 + y_n^2)}{y_n},$$

these properties of $\{z_n\}$ are equivalent to results on Diophantine approximations to ζ. We give this investigation of $\{z_n\}$ in the first part, and the applications in the second part, of the paper.

It is quite obvious that in a similar way matrices of any order may be considered; we then shall get results on the simultaneous approximations of a system of P-adic numbers. I intend to study this problem and the analogous one in the real field in a later paper.

Part I: The Representative of a P-adic Integer

1. The sequences $Z(\zeta)$ and $z(\zeta)$

Let Γ be the modular group of all substitutions

$$(1): \qquad Z = \frac{rz + r'}{qz + q'} \qquad (rq' - r'q = 1)$$

with integer coefficients and determinant 1. Two points Z and z in the complex upper half-plane H, which are related by (1) are called equivalent. As is proved in the theory of Γ, to every point Z in H there is exactly one equivalent point z in the domain F of all points $z = x + yi$, for which

$$(2): \quad -\tfrac{1}{2} \leqq x < \tfrac{1}{2}, \qquad x^2 + y^2 > 1, \quad \text{or} \quad -\tfrac{1}{2} \leqq x \leqq 0, \qquad x^2 + y^2 = 1;$$

this domain F is called the fundamental domain of Γ.—It is useful to define the *distance* between two points z_1 and z_2 in F by

$$\rho(z_1, z_2) = \min\left(| z_1 - z_2 |, | z_1 - z_2 + 1 |, | z_1 - z_2 - 1 |, \left| z_1 + \frac{1}{z_2} \right|, \left| \frac{1}{z_1} + z_2 \right| \right);$$

then the neighborhood of a point z_1 in F consists of all points z_2 in F, which are sufficiently near either to z_1 or one of the equivalent points $z_1 + 1$, $z_1 - 1$, or

$-1/z_1$. In the following considerations, λ denotes a number in F, which is fixed, but arbitrary.

Let P be a natural prime number and ζ a P-adic integer:

$$|\zeta|_P \leqq 1.$$

For every $n = 0, 1, 2, \cdots$ we define two integers A_n and a_n by

(3): $$|\zeta - A_n|_P \leqq P^{-n}, \qquad 0 \leqq A_n \leqq P^n - 1 \qquad (A_0 = 0),$$

and

(4): $$a_n = \frac{A_{n+1} - A_n}{P^n}, \quad \text{so that} \quad 0 \leqq a_n \leqq P - 1 \qquad (a_0 = A_1),$$

and

$$\zeta = \lim_{n\to\infty} A_n = a_0 + a_1 P + a_2 P^2 + \cdots.$$

Therefore ζ is known if the sequence $A(\zeta)$ of all A_n or the sequence $a(\zeta)$ of all a_n are given.

By means of $A(\zeta)$, the sequence $Z(\zeta)$ of all numbers

(5): $$Z_n = \frac{A_n + \lambda}{P^n} \qquad (Z_0 = \lambda)$$

in H is determined; obviously

$$Z_{n+1} = \frac{Z_n + a_n}{P}.$$

A second sequence $z(\zeta)$ of complex numbers z_n is obtained, if by z_n ($n = 0, 1, 2, \cdots$) we understand the number in F which is equivalent to Z_n. This sequence $z(\zeta)$ is called the representative of ζ.

2. The matrices T_n and Ω_n

Let

(6): $$Z_n = \frac{r_n z_n + r_n'}{q_n z_n + q_n'} \quad \text{or} \quad z_n = \frac{-q_n' Z_n + r_n'}{q_n Z_n - r_n} \qquad (r_n q_n' - q_n r_n' = 1)$$

be the modular substitution which connects Z_n with z_n, and put

(7): $$p_n = P^n r_n - A_n q_n, \qquad p_n' = P^n r_n' - A_n q_n'.$$

Then from (5)

(8): $$\lambda = \frac{p_n z_n + p_n'}{q_n z_n + q_n'} \quad \text{or} \quad z_n = \frac{-q_n' \lambda + p_n'}{q_n \lambda - p_n},$$

or symbolically

$$\lambda = \mathrm{T}_n z_n, \qquad z_n = \mathrm{T}_n^{-1} \lambda,$$

where $\mathbf{T}_n$ denotes the matrix

(9): $$\mathbf{T}_n = \begin{pmatrix} p_n & p_n' \\ q_n & q_n' \end{pmatrix}.$$

Since $r_n q_n' - r_n' q_n = 1$, $\mathbf{T}_n$ has the determinant

(10): $$p_n q_n' - p_n' q_n = P^n,$$

and its two lower elements q_n and q_n' are relatively prime. At least one of the q's, therefore, is not divisible by P, say q_n^*; by p_n^* we denote the element of $\mathbf{T}_n$ above q_n^*. In particular for $n = 0$,

$$\mathbf{T}_0 = \mathbf{E} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad p_0^* = 0, \qquad q_0^* = 1,$$

since $Z_0 = \lambda$ lies in F and therefore $Z_0 = z_0$.

The sequence $\mathbf{T}(\zeta)$ of all matrices $\mathbf{T}_n$ has a number of simple properties. Obviously

$$\mathbf{T}_n = \begin{pmatrix} P^n & -A_n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r_n & r_n' \\ q_n & q_n' \end{pmatrix}.$$

Therefore, if for $n = 0, 1, 2, \cdots$

(11): $$\Omega_{n+1} = \mathbf{T}_n^{-1} \mathbf{T}_{n+1} = \begin{pmatrix} \alpha_{n+1} & \alpha_{n+1}' \\ \beta_{n+1} & \beta_{n+1}' \end{pmatrix},$$

say, then Ω_{n+1} has the determinant

(12): $$\alpha_{n+1}\beta_{n+1}' - \alpha_{n+1}'\beta_{n+1} = P.$$

Also

$$\Omega_{n+1} = \begin{pmatrix} r_n & r_n' \\ q_n & q_n' \end{pmatrix}^{-1} \begin{pmatrix} P^n & -A_n \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} P^{n+1} & -A_{n+1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r_{n+1} & r_{n+1}' \\ q_{n+1} & q_{n+1}' \end{pmatrix},$$

or after a simplification

$$\Omega_{n+1} = -\begin{pmatrix} r_n & r_n' \\ q_n & q_n' \end{pmatrix}^{-1} \begin{pmatrix} P & -a_n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r_{n+1} & r_{n+1}' \\ q_{n+1} & q_{n+1}' \end{pmatrix},$$

so that all elements of Ω_{n+1} are integers. We denote the sequence of all Ω_{n+1} by $\Omega(\zeta)$. It can be determined from $\mathbf{T}(\zeta)$ by means of (11); conversely

(13): $$\mathbf{T}_0 = \mathbf{E}, \qquad \mathbf{T}_n = \Omega_1 \Omega_2 \cdots \Omega_n \quad \text{for} \quad n = 1, 2, 3, \cdots,$$

so that $\mathbf{T}(\zeta)$ follows also from $\Omega(\zeta)$. Obviously from (8) and (11)

(14): $$z_n = \Omega_{n+1} z_{n+1} = \frac{\alpha_{n+1} z_{n+1} + \alpha_{n+1}'}{\beta_{n+1} z_{n+1} + \beta_{n+1}'}, \qquad z_{n+1} = \Omega_{n+1}^{-1} z_n = \frac{-\beta_{n+1}' z_n + \alpha_{n+1}'}{\beta_{n+1} z_n - \alpha_{n+1}}.$$

Hence all Ω_{n+1} have the following important property: If $\Omega_{n+1}F$ denotes the domain of all points $z' = \Omega_{n+1}z$, where z lies in F, then F and $\Omega_{n+1}F$ have at least one common point.

3. The theorem of finiteness

From the last remark, we deduce the result:

THEOREM 1: *All terms Ω_{n+1} of the sequence $\Omega(\zeta)$ belong to a finite set $M(P)$ of integer matrices of determinant P, which depends only on P.*

PROOF: Since it is just as easy, we prove a more general result. Suppose that Q is an arbitrary positive integer; let $M(Q)$ denote the set of all integer matrices Ω of determinant Q, for which the two domains F and ΩF have at least one common point. Then we prove that $M(Q)$ is a finite set. Now every integer matrix Ω of determinant Q can be written as

$$\Omega = \mathrm{P}\Sigma,$$

where P is an element of Γ, and

$$\Sigma = \begin{pmatrix} u & v \\ 0 & w \end{pmatrix}$$

is an integer matrix, such that

$$uw = Q, \qquad u > 0, \qquad w > 0, \qquad 0 \leqq v \leqq w - 1.$$

Obviously, there are only a finite number of matrices Σ with these properties, say the matrices

$$\Sigma_h \qquad (h = 1, 2, \cdots, \sigma).$$

The transformed domains

$$\Sigma_h F$$

all lie entirely in the part of the z-plane given by

$$|x| \leqq Q, \qquad y \geqq \frac{\sqrt{3}}{2Q},$$

and therefore enter only a finite number of triangles of the modular division of the z-plane. Hence to every Σ_h there can be only a finite number of elements

$$\mathrm{P}_{hk} \qquad (k = 1, 2, \cdots, \rho_h)$$

of Γ, such that $\mathrm{P}_{hk}\Sigma_h$ belongs to $M(Q)$, and the theorem follows at once.

4. The cases $P = 2$, $P = 3$ and $P = 5$ of Theorem 1

The last proof gives a method for the actual determination of $M(Q)$ and in particular of $M(P)$. We give here the results for the smallest prime numbers P.

For this purpose we divide $M(P)$ into three subsets $M_1(P)$, $M_2(P)$ and $M_3(P)$.

An element Ω of $M(P)$ belongs to the first, second or third of these sets, according as the set of the common points of F and ΩF has positive area, or forms an arc of a curve, or consists of a single point. For $P = 2$ both $M_2(P)$ and $M_3(P)$ are empty, but this is not true for any larger prime number P.

In the following tables, the elements of $M_j(P)$ $(j = 1, 2, 3)$ are arranged in such a way that any two of them stand one above the other, if their product is equal to $\mp P\mathbf{E}$, and that they stand alone, if their square has this value. With Ω also $-\Omega$ belongs to $M_j(P)$; these two matrices lead to the same relation between points of F and therefore are not essentially different.

CASE $P = 2$:

$M(2) = M_1(2)$ has 26 elements:

$$\mp\begin{pmatrix} 0 & 2 \\ -1 & 0 \end{pmatrix}, \quad \mp\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mp\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}, \quad \mp\begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}, \quad \mp\begin{pmatrix} 0 & 2 \\ -1 & 1 \end{pmatrix},$$

$$\mp\begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}, \quad \mp\begin{pmatrix} -1 & 1 \\ 0 & -2 \end{pmatrix}, \quad \mp\begin{pmatrix} -1 & -1 \\ 0 & -2 \end{pmatrix}, \quad \mp\begin{pmatrix} -1 & 2 \\ -1 & 0 \end{pmatrix},$$

$$\mp\begin{pmatrix} 0 & 2 \\ -1 & -1 \end{pmatrix}, \quad \mp\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix},$$

$$\mp\begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}, \quad \mp\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}.$$

CASE $P = 3$:

$M_1(3)$ has 26 elements:

$$\mp\begin{pmatrix} 0 & 3 \\ -1 & 0 \end{pmatrix}, \quad \mp\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mp\begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}, \quad \mp\begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix}, \quad \mp\begin{pmatrix} 0 & 3 \\ -1 & 1 \end{pmatrix},$$

$$\mp\begin{pmatrix} -1 & 0 \\ 0 & -3 \end{pmatrix}, \quad \mp\begin{pmatrix} -1 & 1 \\ 0 & -3 \end{pmatrix}, \quad \mp\begin{pmatrix} -1 & -1 \\ 0 & -3 \end{pmatrix}, \quad \mp\begin{pmatrix} -1 & 3 \\ -1 & 0 \end{pmatrix},$$

$$\mp\begin{pmatrix} 0 & 3 \\ -1 & -1 \end{pmatrix}, \quad \mp\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix},$$

$$\mp\begin{pmatrix} 1 & 3 \\ -1 & 0 \end{pmatrix}, \quad \mp\begin{pmatrix} -1 & 2 \\ -1 & -1 \end{pmatrix};$$

$M_2(3)$ has 4 elements:

$$\mp\begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix},$$

$$\mp\begin{pmatrix} -1 & 1 \\ -1 & -2 \end{pmatrix};$$

$M_3(3)$ has 6 elements:

$$\mp\begin{pmatrix}-1 & -2\\ 2 & 1\end{pmatrix},\quad \mp\begin{pmatrix}0 & -3\\ 1 & 2\end{pmatrix},\quad \mp\begin{pmatrix}-2 & -3\\ 1 & 0\end{pmatrix}.$$

CASE $P = 5$:

$M_1(5)$ has 58 elements:

$$\mp\begin{pmatrix}0 & 5\\ -1 & 0\end{pmatrix},\ \mp\begin{pmatrix}5 & 0\\ 0 & 1\end{pmatrix},\ \mp\begin{pmatrix}5 & 1\\ 0 & 1\end{pmatrix},\ \mp\begin{pmatrix}5 & -1\\ 0 & 1\end{pmatrix},\ \mp\begin{pmatrix}5 & 2\\ 0 & 1\end{pmatrix},$$

$$\mp\begin{pmatrix}-1 & 0\\ 0 & -5\end{pmatrix},\ \mp\begin{pmatrix}-1 & 1\\ 0 & -5\end{pmatrix},\ \mp\begin{pmatrix}-1 & -1\\ 0 & -5\end{pmatrix},\ \mp\begin{pmatrix}-1 & 2\\ 0 & -5\end{pmatrix},$$

$$\mp\begin{pmatrix}5 & -2\\ 0 & 1\end{pmatrix},\ \mp\begin{pmatrix}1 & -5\\ 1 & 0\end{pmatrix},\ \mp\begin{pmatrix}2 & -5\\ 1 & 0\end{pmatrix},\ \mp\begin{pmatrix}0 & -5\\ 1 & 1\end{pmatrix},$$

$$\mp\begin{pmatrix}-1 & -2\\ 0 & -5\end{pmatrix},\ \mp\begin{pmatrix}0 & -5\\ 1 & -1\end{pmatrix},\ \mp\begin{pmatrix}0 & -5\\ 1 & -2\end{pmatrix},\ \mp\begin{pmatrix}-1 & -5\\ 1 & 0\end{pmatrix},$$

$$\mp\begin{pmatrix}0 & -5\\ 1 & 2\end{pmatrix},\ \mp\begin{pmatrix}1 & -4\\ 1 & 1\end{pmatrix},\ \mp\begin{pmatrix}1 & -3\\ 1 & 2\end{pmatrix},\ \mp\begin{pmatrix}1 & -2\\ 2 & 1\end{pmatrix},$$

$$\mp\begin{pmatrix}-2 & -5\\ 1 & 0\end{pmatrix},\ \mp\begin{pmatrix}-1 & -4\\ 1 & -1\end{pmatrix},\ \mp\begin{pmatrix}-2 & -3\\ 1 & -1\end{pmatrix},\ \mp\begin{pmatrix}-1 & -2\\ 2 & -1\end{pmatrix},$$

$$\mp\begin{pmatrix}2 & -3\\ 1 & 1\end{pmatrix},\ \mp\begin{pmatrix}2 & -1\\ 1 & 2\end{pmatrix},$$

$$\mp\begin{pmatrix}-1 & -3\\ 1 & -2\end{pmatrix},\ \mp\begin{pmatrix}-2 & -1\\ 1 & -2\end{pmatrix};$$

$M_2(5)$ has 6 elements:

$$\mp\begin{pmatrix}1 & 3\\ -2 & -1\end{pmatrix},\quad \mp\begin{pmatrix}1 & -2\\ 1 & 3\end{pmatrix},\quad \mp\begin{pmatrix}-3 & -2\\ 1 & -1\end{pmatrix};$$

$M_3(5)$ has 4 elements:

$$\mp\begin{pmatrix}0 & -5\\ 1 & 3\end{pmatrix},\quad \mp\begin{pmatrix}-3 & -5\\ 1 & 0\end{pmatrix}.$$

$M_1(7)$ has 90, $M_2(7)$ has 6, $M_3(7)$ has 10, $M_1(11)$ has 190 elements, etc.

5. The first existence theorem

In this and the following paragraphs we shall prove some theorems, which show that every element of $M(P)$ actually occurs in the set $\Omega(\zeta)$ of a certain P-adic number ζ; if the element lies in $M_2(P)$ or $M_3(P)$, then it may, however, be necessary to choose λ appropriately.

THEOREM 2: *Let*

$$\{\Omega_{n+1}\} = \{\Omega_1, \Omega_2, \Omega_3, \cdots\}, \quad \Omega_{n+1} = \begin{pmatrix} \alpha_{n+1}, & \alpha'_{n+1} \\ \beta_{n+1}, & \beta'_{n+1} \end{pmatrix},$$

be an infinite sequence of integer matrices of determinant P, and define a second sequence of matrices

$$\{\mathbf{T}_n\} = \{\mathbf{T}_0, \mathbf{T}_1, \mathbf{T}_2, \cdots\}$$

by

$$\mathbf{T}_0 = \mathbf{E}, \qquad \mathbf{T}_n = \Omega_1\Omega_2 \cdots \Omega_n = \begin{pmatrix} p_n & p'_n \\ q_n & q'_n \end{pmatrix} \quad \text{for} \quad n = 1, 2, 3, \cdots.$$

Then there is a P-adic integer ζ such that $\{\Omega_{n+1}\} = \Omega(\zeta)$ and $\{\mathbf{T}_n\} = \mathbf{T}(\zeta)$, if and only if $(q_n, q'_n) = 1$ and $\dfrac{-q'_n\lambda + p'_n}{q_n\lambda - p_n}$ lies in F for $n = 0, 1, 2, \cdots$.

PROOF: That the conditions are necessary, was already shown in §2; therefore we have only to prove that they are sufficient.

As in §2, let q_n^* be that one of the two numbers q_n and q'_n which is not divisible by P, and p_n^* the corresponding other element of $\mathbf{T}_n$. Then there is exactly one integer A_n such that

$$p_n^* + q_n^* A_n \equiv 0 \pmod{P^n}, \qquad 0 \leqq A_n \leqq P^n - 1,$$

and since $P \nmid q_n^*$ and $p_n q'_n - p'_n q_n = P^n$, obviously

$$p_n + q_n A_n \equiv 0 \pmod{P^n}, \qquad p'_n + q'_n A_n \equiv 0 \pmod{P^n}.$$

Therefore there are two integers r_n and r'_n such that

$$p_n = P^n r_n - A_n q_n, \qquad p'_n = P^n r'_n - A_n q'_n,$$

and the determinant $r_n q'_n - r'_n q_n = 1$. Thus if

$$Z_n = \frac{A_n + \lambda}{P^n} \quad \text{and} \quad z_n = \frac{-q'_n Z_n + r'_n}{q_n Z_n - r_n},$$

then

$$z_n = \frac{-q'_n\lambda + p'_n}{q_n\lambda - p_n}$$

lies in F by hypothesis.

I assert that

$$A_{n+1} \equiv A_n \pmod{P^n}.$$

For

$$\mathbf{T}_{n+1} = \mathbf{T}_n \Omega_{n+1} = \begin{pmatrix} p_n\alpha_{n+1} + p'_n\beta_{n+1}, & p_n\alpha'_{n+1} + p'_n\beta'_{n+1} \\ q_n\alpha_{n+1} + q'_n\beta_{n+1}, & q_n\alpha'_{n+1} + q'_n\beta'_{n+1} \end{pmatrix},$$

hence

$$p_{n+1} + q_{n+1}A_n \equiv p_{n+1} + q_{n+1}A_{n+1} \equiv p'_{n+1} + q'_{n+1}A_n \equiv p'_{n+1} + q'_{n+1}A_{n+1} \equiv 0 \pmod{P^n},$$

from which the congruence follows immediately. If ζ is defined by

$$\zeta = \lim_{n\to\infty} A_n,$$

then this limit exists as a P-adic integer and satisfies all the conditions.

COROLLARY: *Let* $\mathbf{T} = \begin{pmatrix} p & p' \\ q & q' \end{pmatrix}$ *be an integer matrix of determinant* P^n, *such that* $(q, q') = 1$, *and* $\dfrac{-q'\lambda + p'}{q\lambda - p}$ *lies in* F. *Then there is a* P*-adic integer* ζ, *such that* $\mathbf{T}$ *is the element of* $\mathbf{T}(\zeta)$ *with index* n.

PROOF: We define the integer A_n by

$$p + A_n q \equiv p' + A_n q' \equiv 0 \pmod{P}, \qquad 0 \leqq A_n \leqq P^n - 1,$$

and take a P-adic integer ζ, for which $|\zeta - A_n|_P \leqq P^{-n}$; then by the proof of the last theorem, ζ satisfies all conditions.

6. The second existence theorem

THEOREM 3: *Let* z^* *be an arbitrary number in* F, α *and* β *two relatively prime integers,* ϵ *an arbitrarily small positive number. Then there is an index* $n > 0$ *and a* P*-adic integer* ζ, *such that the elements* $\mathbf{T}_n = \begin{pmatrix} p_n & p'_n \\ q_n & q'_n \end{pmatrix}$ *of* $\mathbf{T}(\zeta)$ *and* z_n *of* $z(\zeta)$ *satisfy the conditions*

$$\alpha q_n + \beta q'_n \equiv \mp 1 \pmod{P}, \qquad |z_n - z^*| \leqq \epsilon.$$

PROOF: Since $(\alpha, \beta) = 1$, there is an integer unimodular matrix

$$\Omega = \begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \end{pmatrix}.$$

Let D be the set of all points z in the upper half plane H, for which

$$\Omega z = \frac{\alpha z + \alpha'}{\beta z + \beta'}$$

satisfies the conditions

$$\Omega z \text{ lies in } F, \qquad |\Omega z - z^*| \leqq \epsilon.$$

Obviously this set has at least one inner point, say the point μ.

For $\eta = \mp 1$, consider the matrix $\Phi_\eta = \begin{pmatrix} 0 & -P \\ 1 & \eta \end{pmatrix}$ of determinant P, and the corresponding substitution

$$z' = \frac{-P}{z+\eta} = \Phi_\eta z.$$

Obviously, this can be written as

$$\frac{2z' + \eta - (1-4P)^{\frac{1}{2}}}{2z' + \eta + (1-4P)^{\frac{1}{2}}} = \vartheta \frac{2z + \eta - (1-4P)^{\frac{1}{2}}}{2z + \eta + (1-4P)^{\frac{1}{2}}},$$

where the factor

$$\vartheta = \frac{\eta - (1-4P)^{\frac{1}{2}}}{\eta + (1-4P)^{\frac{1}{2}}}$$

is a complex number of modulus 1. Since

$$\vartheta^2 + \left(2 - \frac{1}{P}\right)\vartheta + 1 = 0,$$

ϑ is not an algebraic integer, and therefore not a root of unity. Hence the powers $\vartheta^{-1}, \vartheta^{-2}, \vartheta^{-3}, \cdots$ lie everywhere dense on the unit circle in the complex plane.

Therefore, if Z is an arbitrary point in the upper half-plane H, and $K_\eta(Z)$ the circle

$$\left|\frac{2z + \eta - (1-4P)^{\frac{1}{2}}}{2z + \eta + (1-4P)^{\frac{1}{2}}}\right| = \left|\frac{2Z + \eta - (1-4P)^{\frac{1}{2}}}{2Z + \eta + (1-4P)^{\frac{1}{2}}}\right|$$

through this point, then with z also $\Phi_\eta z$ lies on $K_\eta(Z)$, and the points

$$z, \quad \Phi_\eta^{-1} z, \quad \Phi_\eta^{-2} z, \quad \Phi_\eta^{-3} z, \quad \cdots \qquad \left(\Phi_\eta^{-1} z = \frac{-\eta z - P}{z}\right)$$

obtained from z by repeated application of Φ_η^{-1} lie everywhere dense on this circle.

Now it is easy to see that we can connect λ with μ by means of a continuous curve C, which consists of arcs of the two circles $K_{-1}(\lambda)$ and $K_{-1}(\mu)$ and of circles $K_\eta(k + \frac{1}{2} + (1-4P)^{\frac{1}{2}})$, where $\eta = \mp 1$ and $k = 0, 1, 2, \cdots$; for any two circles $K_{-1}(k + \frac{1}{2} + (1-4P)^{\frac{1}{2}})$ and $K_{+1}(\kappa + \frac{1}{2} + (1-4P)^{\frac{1}{2}})$, where $k = 0, 1, 2, \cdots$ and $k - \frac{1}{2} \leqq \kappa \leqq k + 3/2$, intersect. Let $\lambda, \lambda_1, \lambda_2, \cdots, \lambda_t, \mu$ be the successive end points of the arcs composing C. Then by alternate applications of powers of Φ_{-1}^{-1} and Φ_{+1}^{-1}, we can transform λ successively into points arbitrarily near to $\lambda_1, \lambda_2, \cdots, \lambda_t$, and finally into a point z^{**} near to μ, which lies in D. In this way, we obtain an integer matrix

$$T^* = \begin{pmatrix} p^* & p^{*\prime} \\ q^* & q^{*\prime} \end{pmatrix} = \Phi_{+1}^{h_1} \Phi_{-1}^{k_1} \Phi_{+1}^{h_2} \cdots \Phi_{-1}^{k_u},$$

where the exponents h, k are integers $\geqq 0$, such that

$$\lambda = \mathrm{T}^* z^{**}, \quad \text{or} \quad z^{**} = \mathrm{T}^{*-1}\lambda;$$

the determinant of T^* is a power of P, say P^n. Since

$$\Phi_\eta \equiv \begin{pmatrix} 0 & 0 \\ 1 & \eta \end{pmatrix} \pmod{P},$$

necessarily

$$\mathrm{T}^* \equiv \begin{pmatrix} 0 & 0 \\ \mp 1 & \mp 1 \end{pmatrix} \pmod{P}.$$

Now put

$$z_n = \Omega z^{**}, \qquad \mathrm{T}_n = \mathrm{T}^*\Omega^{-1} = \begin{pmatrix} p_n & p_n' \\ q_n & q_n' \end{pmatrix},$$

so that

$$\lambda = \mathrm{T}_n z_n \quad \text{and} \quad z_n = \mathrm{T}_n^{-1}\lambda,$$

and, by the definition of D,

$$z_n \text{ lies in } F, \quad \text{and} \quad |z_n - z^*| \leqq \epsilon.$$

We have

$$\mathrm{T}^* = \mathrm{T}_n\Omega = \begin{pmatrix} p_n & p_n' \\ q_n & q_n' \end{pmatrix}\begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \end{pmatrix} \equiv \begin{pmatrix} 0 & 0 \\ \mp 1 & \mp 1 \end{pmatrix} \pmod{P};$$

therefore

$$\alpha q_n + \beta q_n' \equiv \mp 1 \pmod{P},$$

and from $p_n q_n' - q_n p_n' = P^n$ in particular $(q_n, q_n') = 1$. Hence, by the corollary to Theorem 2, the theorem follows immediately.

7. An application of Theorem 3

To every element Ω of $M(P)$ belongs a certain set $S(\Omega)$ of points z in F, such that the transformed points $\Omega^{-1}z$ also lie in F; this set has inner points, or consists of an arc of a curve, or of a single point, according as Ω belongs to $M_1(P)$, or to $M_2(P)$, or to $M_3(P)$.

THEOREM 4: *Let Ω be an element of $M_1(P)$. Then there is a P-adic integer ζ, such that Ω is an element of the sequence $\Omega(\zeta)$. Provided that λ is chosen suitably, this is also true, if Ω belongs to $M_2(P)$ or $M_3(P)$.*

PROOF: Suppose that z^* is a point of $S(\Omega)$, in particular an inner point, if Ω belongs to $M_1(P)$. If $\Omega = \begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \end{pmatrix}$, then either $(\alpha, \beta) = 1$ or $(\alpha', \beta') = 1$, for $\alpha\beta' - \alpha'\beta = P$. Denote the pair, which is relatively prime, by α^*, β^*. Then,

by the last theorem, we can find a P-adic integer ζ_0 for which z_n is arbitrarily near to z^*, and for which $\mathbf{T}_n = \begin{pmatrix} p_n & p_n' \\ q_n & q_n' \end{pmatrix}$ satisfies the condition

$$\alpha^* q_n + \beta^* q_n' \equiv \mp 1 \not\equiv 0 \pmod{P}.$$

Thus, in particular, if

$$\mathbf{T}_{n+1} = \begin{pmatrix} p_{n+1} & p_{n+1}' \\ q_{n+1} & q_{n+1}' \end{pmatrix} = \mathbf{T}_n \Omega,$$

then $(q_{n+1}, q_{n+1}') = 1$, since one of the q's is not divisible by P.

Now, if Ω lies in $M_1(P)$, then we can choose ζ_0 in such a way that z_n lies in $S(\Omega)$. If, however, Ω is an element of $M_2(P)$ or $M_3(P)$, then by a slight change of λ, so that z_n remains in F, we can again obtain a ζ_0 for which z_n lies in $S(\Omega)$. Hence in both cases the number

$$z_{n+1} = \Omega^{-1} z_n = \mathbf{T}_{n+1}^{-1} \lambda$$

lies in F. Therefore by the corollary to Theorem 2 there is another P-adic integer ζ, such that Ω and z_{n+1} are the elements of $\Omega(\zeta)$ and $z(\zeta)$ with index $n+1$. The proof of Theorem 2 shows that necessarily $|\zeta - \zeta_0|_P \leqq P^{-n}$; hence $\mathbf{T}_n$ also belongs to $\mathbf{T}(\zeta)$, and Ω is an element of $\Omega(\zeta)$.

8. The sets $m(P)$ and $m'(P)$

Let $\Omega = \begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \end{pmatrix}$ be an element of $M(P)$. Then the fix points of the substitution

$$z' = \frac{\alpha z + \alpha'}{\beta z + \beta'} = \Omega z$$

satisfy the quadratic equation

$$\beta z^2 + (\beta' - \alpha) z - \alpha' = 0,$$

and are therefore

$$(15): \qquad \frac{\alpha - \beta' \mp ((\alpha - \beta')^2 + 4\alpha'\beta)^{\frac{1}{2}}}{2\beta} = \frac{\alpha - \beta' \mp ((\alpha + \beta')^2 - 4P)^{\frac{1}{2}}}{2\beta},$$

if $\beta \neq 0$. Let us suppose that $\beta > 0$, and that both roots (15) are conjugate complex numbers, i.e., that $z' = \Omega z$ is an elliptic substitution. Then if

$$(16): \qquad f_1 = \frac{\alpha - \beta' + ((\alpha + \beta')^2 - 4P)^{\frac{1}{2}}}{2\beta}, \qquad f_2 = \frac{\alpha - \beta' - ((\alpha + \beta')^2 - 4P)^{\frac{1}{2}}}{2\beta},$$

f_1 is the fix point with positive imaginary part, and f_2 the conjugate one. Obviously, $z' = \Omega z$ can be written as

$$(17): \qquad \frac{z' - f_1}{z' - f_2} = \vartheta \frac{z - f_1}{z - f_2},$$

where the multiplicator

$$(18): \qquad \vartheta = \frac{\alpha - \beta f_1}{\alpha - \beta f_2} = \frac{\alpha + \beta' - ((\alpha+\beta')^2 - 4P)^{\frac{1}{2}}}{\alpha + \beta' + ((\alpha+\beta')^2 - 4P)^{\frac{1}{2}}}$$

satisfies the quadratic equation

$$(19): \qquad \vartheta^2 + \left(2 - \frac{(\alpha+\beta')^2}{P}\right)\vartheta + 1 = 0.$$

Let us now denote by $m'(P)$ and $m(P)$ the subset of those elements Ω of $M(P)$, which are elliptic, whose fix point f_1 lies in F, and whose multiplicator ϑ is a root of unity, or is not a root of unity, respectively.

For the elements Ω of $m'(P)$, the coefficient

$$2 - \frac{(\alpha+\beta')^2}{P}$$

in (19) obviously must be equal to 0, or to ∓ 1, or to ∓ 2. Hence necessarily $P = 2$ and $\alpha + \beta' = \mp 2$, or $P = 3$ and $\alpha + \beta' = \mp 3$, or P is arbitrary and $\alpha + \beta' = 0$. By the tables in §4, $m'(2)$ has six elements

$$\mp\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad \mp\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}, \quad \mp\begin{pmatrix} 0 & 2 \\ -1 & 0 \end{pmatrix};$$

$m'(3)$ has six elements

$$\mp\begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}, \quad \mp\begin{pmatrix} -1 & 1 \\ -1 & -2 \end{pmatrix}, \quad \mp\begin{pmatrix} 0 & 3 \\ -1 & 0 \end{pmatrix};$$

and $m'(5)$ has four elements

$$\mp\begin{pmatrix} 1 & 3 \\ -2 & -1 \end{pmatrix}, \quad \mp\begin{pmatrix} 0 & 5 \\ -1 & 0 \end{pmatrix}.$$

If $P \geqq 5$ and Ω belongs to $m'(P)$, then $\alpha + \beta' = 0$, so that

$$\Omega = \begin{pmatrix} \alpha & \alpha' \\ \beta & -\alpha \end{pmatrix} \quad \text{and therefore} \quad \Omega^2 = -P\mathrm{E}.$$

Also in the case of $m'(2)$ and $m'(3)$

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}^2 = 2\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}^2 = 2\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

and

$$\begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}^2 = 3\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} -1 & 1 \\ -1 & -2 \end{pmatrix}^2 = 3\begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}.$$

Hence no element Ω of $m'(P)$ can occur twice in succession in the sequence $\Omega(\zeta)$, since otherwise all elements T_n of $\mathrm{T}(\zeta)$ with sufficiently large n would no longer satisfy the condition $(q_n, q'_n) = 1$.

For the elements of $m(P)$,

(20): $$\alpha + \beta' \not\equiv 0 \pmod{P}.$$

For Ω is elliptic, and therefore

$$(\alpha + \beta')^2 - 4P < 0, \quad \text{i.e.} \quad |\alpha + \beta'| < 2\sqrt{P},$$

and so, if (20) were not true,

$$P < 2\sqrt{P} \quad \text{or} \quad P < 4.$$

The tables in §4 show, however, that (20) holds also for $P = 2$ and $P = 3$.

9. A congruence for the powers of a matrix

Let $\Omega = \begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \end{pmatrix}$ be an element of $m(P)$, or more generally any elliptic integer matrix with determinant P^g ($g \geqq 1$), for which (20) is satisfied. Then the equation

(21): $$\begin{vmatrix} \alpha - \varphi & \alpha' \\ \beta & \beta' - \varphi \end{vmatrix} = \varphi^2 - (\alpha + \beta')\varphi + P^g = 0$$

has two different roots φ_1 and φ_2 in the field of the P-adic numbers. For define $((\alpha + \beta')^2 - 4P^g)^{\frac{1}{2}}$ as a P-adic integer by the convergent series

$$((\alpha + \beta')^2 - 4P^g)^{\frac{1}{2}} = (\alpha + \beta')\left\{1 - \frac{4P^g}{(\alpha + \beta')^2}\right\}^{\frac{1}{2}}$$

$$= (\alpha + \beta')\left\{1 - \frac{2P^g}{(\alpha + \beta')^2} - \frac{2^2P^{2g}}{(\alpha + \beta')^4 \cdot 2!} - \frac{1 \cdot 3 \cdot 2^3P^{3g}}{(\alpha + \beta')^6 \cdot 3!} - \cdots\right\};$$

then these roots are given by

(22): $$\varphi_1 = \frac{\alpha + \beta' + ((\alpha + \beta')^2 - 4P^g)^{\frac{1}{2}}}{2}, \quad \varphi_2 = \frac{\alpha + \beta' - ((\alpha + \beta')^2 - 4P^g)^{\frac{1}{2}}}{2}.$$

Obviously

(23): $$\varphi_1 \equiv \alpha + \beta' \not\equiv 0 \pmod{P}, \qquad \varphi_2 \equiv 0 \pmod{P}.$$

It is well known that for every integer $k \geqq 0$

(24): $$\Omega^k = \varphi_1^k \frac{\Omega - \varphi_2 \mathrm{E}}{\varphi_1 - \varphi_2} + \varphi_2^k \frac{\Omega - \varphi_1 \mathrm{E}}{\varphi_2 - \varphi_1}.$$

Hence, from (23),

(25): $$\Omega^k \equiv (\alpha + \beta')^{k-1}\Omega \pmod{P} \qquad \text{for } k \geqq 1.$$

More generally, if $\mathrm{T} = \begin{pmatrix} p & p' \\ q & q' \end{pmatrix}$ is an integer matrix of determinant P^m, put

$$\mathrm{T}^* = \begin{pmatrix} p^* & p^{*\prime} \\ q^* & q^{*\prime} \end{pmatrix} = \mathrm{T}\Omega^k.$$

Then, from (25),

(26): $$\mathrm{T}^* \equiv (\alpha + \beta')^{k-1}\mathrm{T}\Omega \pmod{P} \qquad \text{for } k \geqq 1.$$

Therefore $(q^*, q^{*\prime}) = 1$, *if*

$(\alpha, \beta) = 1$ and $\alpha q + \beta q' \not\equiv 0 \pmod{P}$, or

$$(\alpha', \beta') = 1 \quad \text{and} \quad \alpha' q + \beta' q' \not\equiv 0 \pmod{P}.$$

From the definition of T^*,

$$(\varphi_1 - \varphi_2)p^* = \varphi_1^k\{p(\alpha - \varphi_2) + p'\beta\} - \varphi_2^k\{p(\alpha - \varphi_1) + p'\beta\},$$
$$(\varphi_1 - \varphi_2)q^* = \varphi_1^k\{q(\alpha - \varphi_2) + q'\beta\} - \varphi_2^k\{q(\alpha - \varphi_1) + q'\beta\},$$

and therefore

(27): $$p^* - \frac{p(\alpha - \varphi_2) + p'\beta}{q(\alpha - \varphi_2) + q'\beta}\, q^* = \frac{\beta}{q(\alpha - \varphi_2) + q'\beta}\, \varphi_2^k P^m.$$

This identity shows that $-p^*/q^*$ tends to a P-adic limit for $k \to \infty$, since $|\varphi_2|_P < 1$, and that this limit lies in the quadratic field given by (21).

10. The third existence theorem

In §1, we defined a generalized distance $\rho(z_1, z_2)$ between two points z_1 and z_2 in F. With this definition, the following theorem holds:

THEOREM 5: *Let* $\Omega = \begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \end{pmatrix}$ *be an element of* $m(P)$, f_1 *its fix point in* F. *Then to every* $\epsilon > 0$ *there is a* P*-adic integer* ζ *such that for all sufficiently large* n

$$\rho(z_n, f_1) \leqq \epsilon.$$

PROOF: Let $\delta > 0$ be a constant to be assigned later, and α^*, β^* the pair α, β or α', β', for which $(\alpha^*, \beta^*) = 1$. By Theorem 3, we can determine an index m and a P-adic integer ζ_0 for which $\mathrm{T}_m = \begin{pmatrix} p_m & p_m' \\ q_m & q_m' \end{pmatrix}$ and z_m satisfy the conditions

$$\alpha^* q_m + \beta^* q_m' \not\equiv 0 \pmod{P}, \qquad |z_m - f_1| \leqq \delta.$$

Suppose that

$$\mathrm{T}_m = \Omega_1\Omega_2 \cdots \Omega_m,$$

where $\Omega_1, \Omega_2, \cdots, \Omega_m$ are the m first elements of $\Omega(\zeta_0)$.

Let us form the matrices

$$\mathrm{T}^*_{m+k} = \begin{pmatrix} p^*_{m+k} & p^{*\prime}_{m+k} \\ q^*_{m+k} & q^{*\prime}_{m+k} \end{pmatrix} = \mathrm{T}_m\Omega^k \qquad (k = 1, 2, 3, \cdots),$$

and the numbers

$$z^*_{m+k} = \Omega^{-k} z_m, \qquad (k = 1, 2, 3, \cdots),$$

which lie in H. From (17) in §8, obviously

$$\frac{z^*_{m+k} - f_1}{z^*_{m+k} - f_2} = \vartheta^{-k} \frac{z_m - f_1}{z_m - f_2},$$

where f_1 and f_2 are the fix points and ϑ the multiplicator of Ω. Since Ω belongs to $m(P)$, ϑ is a complex number of modulus 1, which is not a root of unity. Hence the points z^*_{m+k} for $k = 1, 2, 3, \cdots$ lie on a circle K given by

$$\left|\frac{z^* - f_1}{z^* - f_2}\right| = \left|\frac{z_m - f_1}{z_m - f_2}\right|,$$

and lie everywhere dense on K. To every point z^* on K, there is a point z, equivalent with respect to Γ, in F; it is obvious that we can choose δ and therefore also the radius of K so small that all these points satisfy the inequality

$$\rho(z, f_1) \leqq \epsilon.$$

Hence, in particular, if z_{m+k} is the point in F which is equivalent to z^*_{m+k}, then

$$\rho(z_{m+k}, f_1) \leqq \epsilon \qquad (k = 1, 2, 3, \cdots).$$

Now by the definition of equivalence,

$$z^*_{m+k} = \mathrm{P}_{m+k} z_{m+k} \qquad (k = 1, 2, 3, \cdots),$$

where P_{m+k} is an element of Γ. Similarly

$$z^*_{m+k+1} = \mathrm{P}_{m+k+1} z_{m+k+1}.$$

From

$$\Omega z^*_{m+k+1} = z^*_{m+k}$$

and from the two preceding equations,

(28): $$z_{m+k} = \Omega_{m+k+1} z_{m+k+1} \qquad (k = 1, 2, 3, \cdots),$$

where

$$\Omega_{m+k+1} = \mathrm{P}^{-1}_{m+k} \Omega \mathrm{P}_{m+k+1}.$$

If

$$\Omega_{m+1} = \Omega \mathrm{P}_{m+1},$$

then (28) holds also for $k = 0$.

Put

$$\mathrm{T}_n = \begin{pmatrix} p_n & p'_n \\ q_n & q'_n \end{pmatrix} = \mathrm{T}^*_n \mathrm{P}_n = \mathrm{T}_m \Omega^{n-m} \mathrm{P}_n \qquad (n > m),$$

so that

$$\mathrm{T}_n = \Omega_1 \Omega_2 \cdots \Omega_n, \qquad z_n = \mathrm{T}^{-1}_n \lambda \qquad \text{for } n \geqq m.$$

Then

$$(q_n, q'_n) = 1.$$

For if both q_n and q_n' were divisible by P, then the same would be true for q_n^* and $q_n^{*\prime}$ in

$$\mathbf{T}_n^* = \begin{pmatrix} p_n^* & p_n^{*\prime} \\ q_n^* & q_n^{*\prime} \end{pmatrix} = \mathbf{T}_n \mathbf{P}_n^{-1} = \mathbf{T}_m \Omega^{n-m}.$$

But by the last paragraph

$$\mathbf{T}_n^* \equiv (\alpha + \beta')^{n-m-1} \mathbf{T}_m \Omega \pmod{P} \qquad (n \geqq m+1),$$

and therefore by the construction of $\mathbf{T}_m$, either q_n^* or $q_n^{*\prime}$ is not divisible by P.

It is now clear that the sequence of matrices

$$\mathbf{T}_0 = \mathbf{E}, \qquad \mathbf{T}_n = \Omega_1 \Omega_2 \cdots \Omega_n \qquad (n = 1, 2, 3, \cdots)$$

satisfies the conditions of Theorem 2. Hence there is a P-adic integer ζ (for which, by the way, obviously $|\zeta - \zeta_0|_P \leqq P^{-m}$), such that $\{\mathbf{T}_n\} = \mathbf{T}(\zeta)$ and $\{\Omega_n\} = \Omega(\zeta)$. By construction, the elements z_n of $z(\zeta)$ with $n \geqq m+1$ satisfy the required inequality $\rho(z_n, f_1) \leqq \epsilon$.

COROLLARY: *If ζ is the P-adic integer given by Theorem 5, then, for an infinity of indices,*

$$\mathbf{T}_n = \mathbf{T}_m \Omega^{n-m}, \qquad \Omega_n = \Omega.$$

In particular, these two equations are true for all sufficiently large n, if the fix point f_1 of Ω is an inner point of F.

PROOF: Let K be the circle defined in the proof of the preceding theorem. Since f_1 belongs to F, the points of K in F will form an arc of positive length; if, in particular, f_1 is an inner point of F, and ϵ, i.e. δ is sufficiently small, then K lies entirely in F. Hence in the former case an infinity of z_{m+k}^*, in the latter case all z_{m+k}^* with $k \geqq 0$, will lie in F, and the statements follow immediately from the proof of Theorem 5.

REMARK: Since $\mathbf{T}_n = \mathbf{T}_m \Omega^{n-m}$ for an infinity of indices n, formula (27) in §9 shows that ζ is a quadratic irrational P-adic number. Its sequence $\Omega(\zeta)$ will consist only of matrices Ω for large indices, if f_1 is an inner point of F and ϵ is sufficiently small; hence in this case $\Omega(\zeta)$ will be periodic. If, on the other hand, f_1 is not an inner point of F, or ϵ is not small enough, then the points z_{m+k}^* will not all belong to F, and since ϑ is not a root of unity, it is easy to see that $\Omega(\zeta)$ is not periodic for large indices. Hence Lagrange's Theorem on the periodicity of the continued fraction of a real quadratic irrational number has no analogue for P-adic numbers.

11. The function $Y(P)$

For every element Ω of $m(P)$, let $y(\Omega)$ be the imaginary part of the fix point f_1 of Ω in F, and put

(29): $$Y(P) = \min y(\Omega),$$

where the minimum refers to all elements of $m(P)$. Obviously

(30): $$Y(P) \geqq \frac{\sqrt{3}}{2}.$$

For $P \leqq 5$, the tables in §4 give

$$Y(2) = \frac{\sqrt{7}}{2}, \qquad Y(3) = \sqrt{2}, \qquad Y(5) = 1,$$

corresponding to the matrices

$$\begin{pmatrix} 1 & -2 \\ 1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

with fix points

$$\frac{1 \mp \sqrt{-7}}{2}, \qquad \mp\sqrt{-2}, \qquad \mp i,$$

and multiplicators

$$\frac{-3 - \sqrt{-7}}{4}, \qquad \frac{-1 - 2\sqrt{-2}}{3}, \qquad \frac{3 - 4i}{5}.$$

THEOREM 6: $Y(P) = \frac{\sqrt{3}}{2}$, *if and only if* $P \equiv 1 \pmod 6$.

PROOF: We have to find all matrices $\Omega = \begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \end{pmatrix}$ of $m(P)$ with their fix point f_1 at $\rho = \frac{-1 + \sqrt{-3}}{2}$. This condition will be satisfied if

$$\beta\rho^2 + (\beta' - \alpha)\rho - \alpha' = 0,$$

or, since ρ is a root of the irreducible equation $\rho^2 + \rho + 1 = 0$, if

$$\alpha' = -\beta \quad \text{and} \quad \beta' = \alpha + \beta.$$

Therefore

$$\Omega = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha + \beta \end{pmatrix} \quad \text{and} \quad \alpha^2 + \alpha\beta + \beta^2 = \tfrac{1}{4}\{(2\alpha + \beta)^2 + 3\beta^2\} = P.$$

Hence either $P = 3$ or $P \equiv 1 \pmod 6$. If $P = 3$, then necessarily

$$\Omega = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} \quad \text{or} \quad \Omega = \begin{pmatrix} -1 & 1 \\ -1 & -2 \end{pmatrix},$$

and these matrices do not belong to $m(P)$. If, however, $P \equiv 1 \pmod 6$, then β is not divisible by P, so that

$$2 - \frac{(\alpha + \beta')^2}{P} = -2 + \frac{3\beta^2}{P}$$

is not an integer; hence the matrix $\Omega = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha + \beta \end{pmatrix}$ now does belong to $m(P)$ and has ρ as its fix point f_1.

It is very probable that

$$\lim_{P\to\infty} Y(P) = \frac{\sqrt{3}}{2},$$

but so far I have not been able to prove this.[9a]

For $P \leqq 103$, $Y(P)$ is given in the following table:

$P=$	$Y(P)=$	$P=$	$Y(P)=$	$P=$	$Y(P)=$
2	$\sqrt{(7)}/2$	29	$\sqrt{(91)}/10$	67	$\sqrt{(3)}/2$
3	$\sqrt{(2)}$	31	$\sqrt{(3)}/2$	71	$5.\sqrt{(11)}/18$
5	1	37	$\sqrt{(3)}/2$	73	$\sqrt{(3)}/2$
7	$\sqrt{(3)}/2$	41	$\sqrt{(8)}/3$	79	$\sqrt{(3)}/2$
11	$\sqrt{(35)}/6$	43	$\sqrt{(3)}/2$	83	$\sqrt{(74)}/9$
13	$\sqrt{(3)}/2$	47	$\sqrt{(187)}/14$	89	$2.\sqrt{(10)}/7$
17	$\sqrt{(8)}/3$	53	$\sqrt{(91)}/10$	97	$\sqrt{(3)}/2$
19	$\sqrt{(3)}/2$	59	$\sqrt{(55)}/8$	101	$\sqrt{(65)}/9$
23	$\sqrt{(91)}/10$	61	$\sqrt{(3)}/2$	103	$\sqrt{(3)}/2$

12. The function $Y(\zeta)$

Let x_n and y_n be the real and imaginary parts of the n-th element $z_n = x_n + iy_n$ of the representative $z(\zeta)$ of a P-adic integer ζ. Suppose that

(31): $$Y(\zeta) = \limsup_{n\to\infty} y_n;$$

i.e. $Y(\zeta)$ is the greatest number, such that for every $\epsilon > 0$ and an infinity of indices

$$y_n \geqq Y(\zeta) - \epsilon.$$

Obviously

(32): $$Y(\zeta) \geqq \sqrt{(3)}/2.$$

THEOREM 7: *To every positive number ϵ, there is a P-adic integer ζ, for which*

$$Y(\zeta) \leqq Y(P) + \epsilon.$$

PROOF: Let Ω be an element of $m(P)$, such that $Y(P)$ is the imaginary part of its fix point f_1 in F. By Theorem 5, there is a P-adic integer ζ for any given $\delta > 0$, such that

$$\rho(z_n, f_1) = \min\left(|z_n - f_1|, |z_n - f_1 + 1|, |z_n - f_1 - 1|, \left|z_n + \frac{1}{f_1}\right|, \left|\frac{1}{z_n} + f_n\right|\right) \leqq \delta$$

for all sufficiently large n. This implies in particular

$$|y_n - Y(P)| \leqq \epsilon,$$

[9a] Addendum: This equation has now been proved by H. Davenport; see the following paper, these Annals, pp. 59–62.

if δ is chosen sufficiently small. For all points in the domain F':

$$|x| \leqq \tfrac{1}{2}, \qquad |z| \geqq 1,$$

which are equivalent to f_1, obviously have the same imaginary part $Y(P)$ as f_1; therefore the statement follows immediately from the continuity of the function $\rho(z, f_1)$.

Probably the following conjecture holds:

"*For every prime number P, and for every P-adic integer ζ,*

$$Y(\zeta) \geqq Y(P).$$"

For $P \equiv 1 \pmod 6$, thus in particular for $P = 7$, this conjecture is indeed true, as is evident from Theorem 6 and the trivial inequality (32). I shall prove the following three theorems, which show that it also holds for $P = 2$, $P = 3$, and $P = 5$, and therefore for all prime numbers less than 10.

13. The lower bound of $Y(\zeta)$ for $P = 2$, $P = 3$, and $P = 5$

THEOREM 8: *For every index n and for any diadic integer ζ*

$$\max(y_n, y_{n+1}, y_{n+2}) \geqq \frac{\sqrt{7}}{2}.$$

Hence in particular for all diadic integers ζ

$$Y(\zeta) \geqq \frac{\sqrt{7}}{2} = Y(2).$$

THEOREM 9: *For every index n and for any triadic integer ζ*

$$\max(y_n, y_{n+1}, y_{n+2}) \geqq \sqrt{2}.$$

Hence in particular for all triadic integers ζ

$$Y(\zeta) \geqq \sqrt{2} = Y(3).$$

THEOREM 10: *For every index n and for any pentadic integer ζ*

$$\max(y_n, y_{n+1}) \geqq 1.$$

Hence in particular for all pentadic integers ζ

$$Y(\zeta) \geqq 1 = Y(5).$$

The proofs of these three theorems depend on the following simple considerations. In

(33): $$z_n = \Omega_{n+1} z_{n+1},$$

Ω_{n+1} must be an element of $M(P)$, and both $z_n = x_n + iy_n$ and $z_{n+1} = x_{n+1} + iy_{n+1}$ must be points in F and therefore also in the closed domain F':

$$|x| \leqq \tfrac{1}{2}, \qquad |z| \geqq 1.$$

All we have to do is to determine, for every Ω in $M(P)$, the greatest partial domains S_n of points z_n and S_{n+1} of points z_{n+1} in F', which are related by (33) with $\Omega_{n+1} = \Omega$. In certain cases, these domains may degenerate into an arc of a circle or a segment of a straight line, or into a single point; if they contain inner points, then they are mapped conformally one upon the other.

For some of the elements of $M(P)$, one or the other of the domains S_n, S_{n+1} will contain only points, for which the ordinate $y_n(y_{n+1})$ is sufficiently large, e.g. $\geqq \sqrt{7}/2$ for $P = 2$. In other cases, both y_n and y_{n+1} may assume values smaller than the required bound, but not simultaneously, so that one of them always remains sufficiently large. In still other cases, both y_n and y_{n+1} will become too small at the same time. Then it is necessary to consider *two consecutive* (equal or different) elements Ω_{n+1}, Ω_{n+2} of $\Omega(\zeta)$, one of which is equal to Ω, and the three numbers z_n, z_{n+1}, $z_{n+2} = x_{n+2} + iy_{n+2}$ connected by

(34): $$z_n = \Omega_{n+1} z_{n+1}, \quad z_{n+1} = \Omega_{n+2} z_{n+2}.$$

As in the simpler case above, there will be the three greatest domains S'_n, S'_{n+1}, S'_{n+2} of points z_n, z_{n+1}, z_{n+2} in F', which are connected by (34). The necessity for considering these sets of three domains will arise only in the proofs of Theorems 8 and 9. In all instances it will be easy to show that at least one of the numbers y_n, y_{n+1}, y_{n+2} is always sufficiently large. Only those combinations Ω_{n+1}, Ω_{n+2} have to be considered for which the simpler method fails; and when the product $\Omega_{n+1}\Omega_{n+2}$ is divisible by P, then also this combination can be excluded (compare §8).

Matrices Ω_{n+1} and $-\Omega_{n+1}$ lead to the same relations (33) or (34), so that only one of them has to be considered. Also our method is symmetrical in y_n and y_{n+1} in the case of a single matrix, and symmetrical in y_n and y_{n+2} in the case of two matrices. Hence in the first case, if one of the ordinates is sufficiently large for $\Omega_{n+1} = \Omega$, then the same will be true for $\Omega_{n+1} = \mp P\Omega^{-1}$; and in the second case, combinations Ω_{n+1}, Ω_{n+2} lead to the same bounds as $P\Omega_{n+2}^{-1}$, $P\Omega_{n+1}^{-1}$.

We arrange the steps of the proofs in the form of tables, so as to render them clearer.

14. Proof of Theorem 8

The elements

$$\mp\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \qquad \mp\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}$$

of $M(2)$ satisfy the equations

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}^2 = 2\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = 2\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$\begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}^2 = 2\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

Hence at least one of any two consecutive terms of $\Omega(\zeta)$ is different from these four matrices.

For the remaining elements of $M(2)$, the inequality

$$\max(y_n, y_{n+1}) \geqq \frac{\sqrt{7}}{2}$$

is satisfied, as follows from the following table:

Ω_{n+1}	S_n	S_{n+1}	Inequalities for y_n, y_{n+1}	$-P\Omega_{n+1}^{-1}$	Similar results for the matrices
$\begin{pmatrix} 0 & 2 \\ -1 & 0 \end{pmatrix}$			$\max(\lvert z_n\rvert, \lvert z_{n+1}\rvert) \geqq \sqrt{2}$, $\lvert x_n\rvert \leqq \frac{1}{2}$, $\lvert x_{n+1}\rvert \leqq \frac{1}{2}$, hence $\max(y_n, y_{n+1}) \geqq \frac{\sqrt{7}}{2}$		
$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$	$\lvert x_n\rvert \leqq \frac{1}{2}$, $\lvert z_n\rvert \geqq 2$	$\lvert z_{n+1}\rvert \leqq \frac{1}{4}$, $\lvert z_{n+1}\rvert \geqq 1$	$y_n \geqq \frac{\sqrt{15}}{2} > \frac{\sqrt{7}}{2}$	$\begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}$	
$\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$	$0 \leqq x_n \leqq \frac{1}{2}$, $\lvert z_n - 1\rvert \geqq 2$	$-\frac{1}{2} \leqq x_{n+1} \leqq -\frac{1}{4}$, $\lvert z_{n+1}\rvert \geqq 1$	$y_n \geqq \sqrt{3} > \frac{\sqrt{7}}{2}$	$\begin{pmatrix} -1 & 1 \\ 0 & -2 \end{pmatrix}$	$\begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} -1 & -1 \\ 0 & -2 \end{pmatrix}$
$\begin{pmatrix} 0 & -2 \\ -1 & 1 \end{pmatrix}$	$\lvert z_n - 2\rvert \leqq 2$, $\lvert z_n\rvert \geqq 1$, $x_n \leqq \frac{1}{2}$	$\lvert z_{n+1} - 1\rvert \leqq 2$, $\lvert z_{n+1} + 1\rvert \geqq 2$, $x_{n+1} \leqq \frac{1}{2}$	$y_{n+1} \geqq \frac{\sqrt{7}}{2}$	$\begin{pmatrix} -1 & 2 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 2 \\ -1 & -1 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$

15. Proof of Theorem 9

We divide $M(3)$ into three subsets S_1, S_2, and S_3, where S_1 consists of the 6 elements

$$\mp\begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}, \mp\begin{pmatrix} -1 & 1 \\ -1 & -2 \end{pmatrix}, \pm\begin{pmatrix} -1 & -2 \\ 2 & 1 \end{pmatrix},$$

S_2 of the 4 elements

$$\mp\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}, \mp\begin{pmatrix} -1 & 2 \\ -1 & -1 \end{pmatrix},$$

and S_3 of the remaining 26 elements of $M(3)$. The relations

$$\begin{pmatrix}2&1\\-1&1\end{pmatrix}\begin{pmatrix}2&1\\-1&1\end{pmatrix}=3\begin{pmatrix}1&1\\-1&0\end{pmatrix},\qquad \begin{pmatrix}2&1\\-1&1\end{pmatrix}\begin{pmatrix}-1&1\\-1&-2\end{pmatrix}=3\begin{pmatrix}-1&0\\0&-1\end{pmatrix},$$

$$\begin{pmatrix}-1&1\\-1&-2\end{pmatrix}\begin{pmatrix}2&1\\-1&1\end{pmatrix}=3\begin{pmatrix}-1&0\\0&-1\end{pmatrix},\quad \begin{pmatrix}-1&1\\-1&-2\end{pmatrix}\begin{pmatrix}-1&1\\-1&-2\end{pmatrix}=3\begin{pmatrix}0&-1\\1&1\end{pmatrix},$$

$$\begin{pmatrix}-1&-2\\2&1\end{pmatrix}\begin{pmatrix}2&1\\-1&1\end{pmatrix}=3\begin{pmatrix}0&-1\\1&1\end{pmatrix},\quad \begin{pmatrix}-1&-2\\2&1\end{pmatrix}\begin{pmatrix}-1&1\\-1&-2\end{pmatrix}=3\begin{pmatrix}1&1\\-1&0\end{pmatrix},$$

$$\begin{pmatrix}2&1\\-1&1\end{pmatrix}\begin{pmatrix}-1&-2\\2&1\end{pmatrix}=3\begin{pmatrix}0&-1\\1&1\end{pmatrix},$$

$$\begin{pmatrix}-1&1\\-1&-2\end{pmatrix}\begin{pmatrix}-1&-2\\2&1\end{pmatrix}=3\begin{pmatrix}1&1\\-1&0\end{pmatrix},$$

$$\begin{pmatrix}-1&-2\\2&1\end{pmatrix}\begin{pmatrix}-1&-2\\2&1\end{pmatrix}=3\begin{pmatrix}-1&0\\0&-1\end{pmatrix}$$

show that at least one of any two consecutive terms of $\Omega(\zeta)$ does not belong to S_1.

Similarly it is evident from the equations

$$\begin{pmatrix}1&2\\-1&1\end{pmatrix}\begin{pmatrix}2&1\\-1&1\end{pmatrix}=3\begin{pmatrix}0&1\\-1&0\end{pmatrix},\qquad \begin{pmatrix}1&2\\-1&1\end{pmatrix}\begin{pmatrix}-1&1\\-1&-2\end{pmatrix}=3\begin{pmatrix}-1&-1\\0&-1\end{pmatrix},$$

$$\begin{pmatrix}2&1\\-1&1\end{pmatrix}\begin{pmatrix}-1&2\\-1&-1\end{pmatrix}=3\begin{pmatrix}-1&1\\0&-1\end{pmatrix},\quad \begin{pmatrix}-1&1\\-1&-2\end{pmatrix}\begin{pmatrix}-1&2\\-1&-1\end{pmatrix}=3\begin{pmatrix}0&-1\\1&0\end{pmatrix},$$

$$\begin{pmatrix}1&2\\-1&1\end{pmatrix}\begin{pmatrix}-1&-2\\2&1\end{pmatrix}=3\begin{pmatrix}1&0\\1&1\end{pmatrix},$$

$$\begin{pmatrix}-1&-2\\2&1\end{pmatrix}\begin{pmatrix}-1&2\\-1&-1\end{pmatrix}=3\begin{pmatrix}1&0\\-1&1\end{pmatrix},$$

that two consecutive terms of $\Omega(\zeta)$ cannot have the form

$$\Omega_{n+1}=\begin{pmatrix}1&2\\-1&1\end{pmatrix},\qquad \Omega_{n+2}\text{ in }S_1\quad\text{or}\quad \Omega_{n+1}\text{ in }S_1,\qquad \Omega_{n+2}=\begin{pmatrix}-1&2\\-1&-1\end{pmatrix}.$$

In the twelve remaining cases, we obtain the table shown on page 34.

Next, it is obvious from

$$\begin{pmatrix}1&2\\-1&1\end{pmatrix}\begin{pmatrix}-1&2\\-1&-1\end{pmatrix}=\begin{pmatrix}-1&2\\-1&-1\end{pmatrix}\begin{pmatrix}1&2\\-1&1\end{pmatrix}=3\begin{pmatrix}-1&0\\0&-1\end{pmatrix},$$

that two consecutive elements of $\Omega(\zeta)$ cannot be

$$\Omega_{n+1}=\begin{pmatrix}1&2\\-1&1\end{pmatrix},\qquad \Omega_{n+2}=\begin{pmatrix}-1&2\\-1&-1\end{pmatrix}$$

Ω_{n+1}	Ω_{n+2}	S'_n	S'_{n+1}	S'_{n+2}	Inequalities for y_n, y_{n+1}, y_{n+2}	Similar results for Ω_{n+1}	Similar results for Ω_{n+2}
$\begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$	$z_n = -\frac{1}{2} + \frac{9}{4l}i,$ where $\frac{\sqrt{3}}{2} \leqq l \leqq \frac{3\sqrt{3}}{2}$	$z_{n+1} = \frac{9 - 4l^2 + 12il}{9 + 4l^2}$	$z_{n+2} = -\frac{1}{2} + il$	$\max(y_n, y_{n+2}) \geqq \frac{3}{2} > \sqrt{2}$	$\begin{pmatrix} -1 & 2 \\ -1 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 1 \\ -1 & -2 \end{pmatrix}$
$\begin{pmatrix} -1 & 1 \\ -1 & -2 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$	$z_n = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$	$z_{n+1} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$	$z_{n+2} = -\frac{1}{2} + \frac{3\sqrt{3}}{2}i$	$y_{n+2} = \frac{3\sqrt{3}}{2} > \sqrt{2}$	$\begin{pmatrix} -1 & 2 \\ -1 & -1 \end{pmatrix}$	$\begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$
$\begin{pmatrix} -1 & -2 \\ 2 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$	$z_n = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$	$z_{n+1} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$	$z_{n+2} = -\frac{1}{2} + \frac{3\sqrt{3}}{2}i$	$y_{n+2} = \frac{3\sqrt{3}}{2} > \sqrt{2}$	$\begin{pmatrix} -1 & 2 \\ -1 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & -2 \\ 2 & 1 \end{pmatrix}$

or

$$\Omega_{n+1} = \begin{pmatrix} -1 & 2 \\ -1 & -1 \end{pmatrix}, \qquad \Omega_{n+2} = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}.$$

It is, however, possible that $\Omega_{n+1} = \Omega_{n+2} = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$, or

$$z_n = \frac{z_{n+1} + 2}{-z_{n+1} + 1} = \frac{-z_{n+2} + 4}{-2z_{n+2} - 1}, \qquad z_{n+1} = \frac{z_{n+2} + 2}{-z_{n+2} + 1}.$$

Then

$$y_n = \frac{9y_{n+2}}{(2x_{n+2} + 1)^2 + 2y_{n+2}^2}, \qquad y_{n+1} = \frac{3y_{n+2}}{(x_{n+2} - 1)^2 + 4y_{n+2}^2}.$$

Suppose that $y_{n+1} \leqq \sqrt{2}$, $y_{n+2} \leqq \sqrt{2}$, so that

$$\frac{3y_{n+2}}{(x_{n+2} - 1)^2 + y_{n+2}^2} \leqq \sqrt{2}, \qquad (x_{n+2} - 1)^2 \geqq y_{n+2}\left(\frac{3}{\sqrt{2}} - y_{n+2}\right).$$

Since $\sqrt{3}/2 \leqq y_{n+2} \leqq \sqrt{2}$, and $y\left(\frac{3}{\sqrt{2}} - y\right)$ assumes its minimum in any interval in one of its end points,

$$(x_{n+2} - 1)^2 \geqq \min\left(\sqrt{2}\left(\frac{3}{\sqrt{2}} - \sqrt{2}\right), \frac{\sqrt{3}}{2}\left(\frac{3}{\sqrt{2}} - \frac{\sqrt{3}}{2}\right)\right) = 1.$$

Therefore $-\frac{1}{2} \leqq x_{n+2} \leqq 0$, and finally

$$y_n \geqq \frac{9y_{n+2}}{1 + 4y_{n+2}^2} = \sqrt{2} + \frac{(4\sqrt{2}\, y_{n+2} - 1)(\sqrt{2} - y_{n+2})}{1 + 4y_{n+2}^2} \geqq \sqrt{2}.$$

A similar proof holds if

$$\Omega_{n+1} = \Omega_{n+2} = \begin{pmatrix} -1 & 2 \\ -1 & -1 \end{pmatrix}.$$

Our discussion has now exhausted all cases in which two consecutive elements of $\Omega(\zeta)$ belong to S_1 or S_2. There remain the elements of S_3; for these the table shown on page 36 is obtained.

16. Proof of Theorem 10

For $P = 5$, it is not necessary to consider *two consecutive* elements of $\Omega(\zeta)$, as we shall find that the theorem holds in its stronger form with only two ordinates.

$M(5)$ has 68 elements (see §4). Of these, only the following eight

$$\mp\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}, \mp\begin{pmatrix} -2 & -1 \\ 1 & -2 \end{pmatrix}, \qquad \mp\begin{pmatrix} -1 & -2 \\ 2 & -1 \end{pmatrix}, \mp\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

require particular examination.

Ω_{n+1}	S_n	S_{n+1}	Inequalities for y_n, y_{n+1}	$-P\Omega_{n+1}^{-1}$	Similar results for the matrices:—
$\begin{pmatrix} 0 & 3 \\ -1 & 0 \end{pmatrix}$			$\max(\lvert z_n\rvert, \lvert z_{n+1}\rvert) \geqq \sqrt{3}$, $\lvert x_n\rvert \leqq \frac{1}{2}$, $\lvert x_{n+1}\rvert \leqq \frac{1}{2}$, hence $\max(y_n, y_{n+1}) \geqq \frac{\sqrt{11}}{2} > \sqrt{2}$		
$\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$	$\lvert x_n\rvert \leqq \frac{1}{2}$, $\lvert z_n\rvert \geqq 3$	$\lvert x_{n+1}\rvert \leqq \frac{1}{2}$, $\lvert z_{n+1}\rvert \geqq 1$	$y_n \geqq \frac{\sqrt{35}}{2} > \sqrt{2}$	$\begin{pmatrix} -1 & 0 \\ 0 & -3 \end{pmatrix}$	
$\begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}$	$\lvert x_n\rvert \leqq \frac{1}{2}$, $\lvert z_n - 1\rvert \geqq 3$	$-\frac{1}{2} \leqq x_{n+1} \leqq -\frac{1}{6}$, $\lvert z_{n+1}\rvert \geqq 1$	$y_n \geqq \frac{3\sqrt{3}}{2} > \sqrt{2}$	$\begin{pmatrix} -1 & 1 \\ 0 & -3 \end{pmatrix}$	$\begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} -1 & -1 \\ 0 & -3 \end{pmatrix}$
$\begin{pmatrix} 0 & 3 \\ -1 & 1 \end{pmatrix}$	$\lvert z_n\rvert \geqq 1$, $\lvert z_n - 3\rvert \leqq 3$, $x_n \leqq \frac{1}{2}$	$\lvert z_{n+1} + 2\rvert \geqq 3$, $\lvert z_{n+1} - 1\rvert \leqq 3$, $x_{n+1} \leqq \frac{1}{2}$	$y_{n+1} \geqq \frac{\sqrt{11}}{2} > \sqrt{2}$	$\begin{pmatrix} -1 & 3 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 3 \\ -1 & -1 \end{pmatrix}$, $\begin{pmatrix} 1 & 3 \\ -1 & 0 \end{pmatrix}$
$\begin{pmatrix} 0 & -3 \\ 1 & 2 \end{pmatrix}$	$z_n = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$	$z_{n+1} = -\frac{1}{2} + \frac{3\sqrt{3}}{2}i$	$y_{n+1} = \frac{3\sqrt{3}}{2} > \sqrt{2}$	$\begin{pmatrix} -2 & -3 \\ 1 & 0 \end{pmatrix}$	

If $\Omega_{n+1} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$, then

$$\lvert z_n\rvert^2 - 1 = 3\frac{\left(x_{n+1} - \frac{4}{3}\right)^2 + y_{n+1}^2 - \left(\frac{5}{3}\right)^2}{(x_{n+1} + 2)^2 + y_{n+1}^2},$$

$$y_n - 1 = -\frac{(x_{n+1}+2)^2 + \left(y_{n+1} - \frac{5}{2}\right)^2 - \left(\frac{5}{2}\right)^2}{(x_{n+1}+2)^2 + y_{n+1}^2}.$$

Now every point z_{n+1} for which

$$0 \leqq x_{n+1} \leqq \tfrac{1}{2}, \qquad |z_{n+1}| \geqq 1, \qquad y_{n+1} < 1,$$

lies inside the circle

$$(x_{n+1} - \tfrac{4}{3})^2 + y_{n+1}^2 = (\tfrac{5}{3})^2,$$

and every point for which

$$-\tfrac{1}{2} \leqq x_{n+1} \leqq 0, \qquad |z_{n+1}| \geqq 1, \qquad y_{n+1} < 1,$$

lies inside the second circle

$$(x_{n+1}+2)^2 + (y_{n+1} - \tfrac{5}{2})^2 = (\tfrac{5}{2})^2.$$

Hence in the first case $|z_n| < 1$, so that z_{n+1} does not belong to S_{n+1} ; in the second case $y_n > 1$. This proves that $\max(y_n, y_{n+1}) \geqq 1$ generally. The same result holds for $\begin{pmatrix} -2 & -1 \\ 1 & -2 \end{pmatrix} = -5\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}^{-1}$.

Similarly, if $\Omega_{n+1} = \begin{pmatrix} -1 & -2 \\ 2 & -1 \end{pmatrix}$, then

$$|z_n|^2 - 1 = -3\frac{\left(x_{n+1} - \frac{4}{3}\right)^2 + y_{n+1}^2 - \left(\frac{5}{3}\right)^2}{(2x_{n+1}-1)^2 + 4y_{n+1}^2},$$

$$y_n - 1 = -4\frac{\left(x_{n+1} - \frac{1}{2}\right)^2 + \left(y_{n+1} - \frac{5}{8}\right)^2 - \left(\frac{5}{8}\right)^2}{(2x_{n+1}-1)^2 + 4y_{n+1}^2}.$$

Hence again the points z_{n+1} for which

$-\tfrac{1}{2} \leqq x_{n+1} \leqq 0$, $|z_{n+1}| \geqq 1$, $y_{n+1} < 1$, i.e. $(x_{n+1} - \tfrac{4}{3})^2 + y_{n+1}^2 - (\tfrac{5}{3})^2 > 0$,

do not belong to S_{n+1}, and $y_n > 1$ for all points z_{n+1} for which

$0 \leqq x_{n+1} \leqq \tfrac{1}{2}$, $|z_{n+1}| \geqq 1$, $y_{n+1} < 1$, i.e. $(x_{n+1} - \tfrac{1}{2})^2 + (y_{n+1} - \tfrac{5}{8})^2 - (\tfrac{5}{8})^2 < 0$.

Thus always $\max(y_n, y_{n+1}) \geqq 1$, and the same result holds for

$$\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = -5\begin{pmatrix} -1 & -2 \\ 2 & -1 \end{pmatrix}^{-1}.$$

For the remaining 60 elements of $M(5)$, the table shown on page 38 is obtained:

Ω_{n+1}	S_n	S_{n+1}	Inequalities for y_n, y_{n+1}	$-5\Omega_{n+1}^{-1}$	Similar results for the matrices:—
$\begin{pmatrix} 0 & 5 \\ -1 & 0 \end{pmatrix}$			$\max(\lvert z_n\rvert, \lvert z_{n+1}\rvert) \geqq \sqrt{5}$, $\lvert x_n\rvert \leqq \frac{1}{2}$, $\lvert x_{n+1}\rvert \leqq \frac{1}{2}$, hence $\max(y_n, y_{n+1}) \geqq \frac{\sqrt{19}}{2} > 1.$		
$\begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}$	$\lvert x_n\rvert \leqq \frac{1}{2}, \lvert z_n\rvert \geqq 5$	$\lvert x_{n+1}\rvert \leqq \frac{1}{10}, \lvert z_{n+1}\rvert \geqq 1$	$y_n \geqq \frac{3\sqrt{11}}{2} > 1$	$\begin{pmatrix} -1 & 0 \\ 0 & -5 \end{pmatrix}$	
$\begin{pmatrix} 5 & 1 \\ 0 & 1 \end{pmatrix}$	$\lvert x_n\rvert \leqq \frac{1}{2}, \lvert z_n - 1\rvert \geqq 5$	$-\frac{3}{10} \leqq x_{n+1} \geqq \frac{1}{10}$, $\lvert z_{n+1}\rvert \geqq 1$	$y_n \geqq \frac{\sqrt{91}}{2} > 1$	$\begin{pmatrix} -1 & 1 \\ 0 & -5 \end{pmatrix}$	$\begin{pmatrix} 5 & -1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & -1 \\ 0 & -5 \end{pmatrix}$
$\begin{pmatrix} 5 & 2 \\ 0 & 1 \end{pmatrix}$	$\lvert x_n\rvert \leqq \frac{1}{2}, \lvert z_n - 2\rvert \geqq 5$	$-\frac{1}{2} \leqq x_{n+1} \geqq -\frac{3}{10}$, $\lvert z_{n+1}\rvert \geqq 1$	$y_n \geqq \frac{5\sqrt{3}}{2} > 1$	$\begin{pmatrix} -1 & 2 \\ 0 & -5 \end{pmatrix}$	$\begin{pmatrix} 5 & -2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & -2 \\ 0 & -5 \end{pmatrix}$
$\begin{pmatrix} 1 & -5 \\ 1 & 0 \end{pmatrix}$	$\lvert x_n\rvert \leqq \frac{1}{2}, \lvert z_n - 1\rvert \leqq 5$, $\lvert z_n + 4\rvert \geqq 5$	$x_{n+1} \leqq \frac{1}{2}, \lvert z_{n+1}\rvert \geqq 1$, $\lvert z_{n+1} - 5\rvert \leqq 5$, $\left\lvert z_{n+1} - \frac{5}{3}\right\rvert \geqq \frac{5}{3}$	$y_n \geqq \frac{\sqrt{19}}{2} > 1$	$\begin{pmatrix} 0 & -5 \\ 1 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & -5 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -5 \\ 1 & 1 \end{pmatrix}$

$\begin{pmatrix} 2 & -5 \\ 1 & 0 \end{pmatrix}$	$x_n \leqq \frac{1}{2}, \lvert z_n - 2 \rvert \leqq 5$ $\lvert z_n + 3 \rvert \geqq 5$	$x_{n+1} \leqq \frac{1}{2}, \lvert z_{n+1} \rvert \geqq 1$ $\left\lvert z_{n+1} - \frac{5}{2} \right\rvert \leqq \frac{5}{2}$	$y_n \geqq \frac{\sqrt{51}}{2} > 1$	$\begin{pmatrix} 0 & -5 \\ 1 & -2 \end{pmatrix}$	$\begin{pmatrix} -2 & -5 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -5 \\ 1 & 2 \end{pmatrix}$
$\begin{pmatrix} 1 & -4 \\ 1 & 1 \end{pmatrix}$	$\lvert x_n \rvert \leqq \frac{1}{2}, \lvert z_n + 4 \rvert \leqq 5$ $\left\lvert z_n + \frac{2}{3} \right\rvert \geqq \frac{5}{3}$	$\lvert x_{n+1} \rvert \leqq \frac{1}{2}, \lvert z_n - 4 \rvert \leqq 5$ $\left\lvert z_{n+1} - \frac{2}{3} \right\rvert \geqq \frac{5}{3}$	$\left.\begin{matrix} y_n \\ y_{n+1} \end{matrix}\right\} \geqq \frac{\sqrt{51}}{6} > 1$	$\begin{pmatrix} -1 & -4 \\ 1 & -1 \end{pmatrix}$	
$\begin{pmatrix} 2 & -3 \\ 1 & 1 \end{pmatrix}$	$\lvert x_n \rvert \leqq \frac{1}{2}, \lvert z_n + 3 \rvert \leqq 5$ $\left\lvert z_n - \frac{1}{3} \right\rvert \geqq \frac{5}{3}$	$\lvert x_{n+1} \rvert \leqq \frac{1}{2}, \lvert z_{n+1} \rvert \geqq 1$ $\left\lvert z_{n+1} - \frac{2}{3} \right\rvert \leqq \frac{5}{3}$	$y_n \geqq \frac{5\sqrt{3}}{6} > 1$	$\begin{pmatrix} -1 & -3 \\ 1 & -2 \end{pmatrix}$	$\begin{pmatrix} -2 & -3 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & -3 \\ 1 & 2 \end{pmatrix}$
$\begin{pmatrix} 1 & 3 \\ -2 & -1 \end{pmatrix}$	$z_n = -\frac{1}{2} + ti,$ where $\frac{\sqrt{3}}{2} \leqq t \leqq \frac{5\sqrt{3}}{6}$	$z_{n+1} = -\frac{1}{2} + \frac{5}{4t} i$	$\max(y_n, y_{n+1})$ $\geqq \frac{\sqrt{5}}{2} > 1$		
$\begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix}$	$z_n = \frac{-25 + 4t^2 + 20it}{25 + 4t^2}$ where $\frac{5\sqrt{3}}{6} \leqq t \leqq \frac{5\sqrt{3}}{2}$	$z_{n+1} = -\frac{1}{2} + ti$	$y_{n+1} \geqq \frac{5\sqrt{3}}{6} > 1$	$\begin{pmatrix} -3 & -2 \\ 1 & -1 \end{pmatrix}$	
$\begin{pmatrix} 0 & -5 \\ 1 & 3 \end{pmatrix}$	$z_n = -\frac{1}{2} + \frac{\sqrt{3}}{2} i$	$z_{n+1} = -\frac{1}{2} + \frac{5\sqrt{3}}{2} i$	$y_{n+1} = \frac{5\sqrt{3}}{2} > 1$	$\begin{pmatrix} -3 & 5 \\ 1 & 0 \end{pmatrix}$	

17. A characteristic property of a rational number

Let us return to the case of a general prime number P, and divide the elements $\Omega = \begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \end{pmatrix}$ of $M(P)$ into two sets $M'(P)$ and $M''(P)$, such that

$$\beta \neq 0 \text{ for the elements of } M'(P),$$

and

$$\beta = 0 \text{ for the elements of } M''(P).$$

THEOREM 11. *Suppose that ζ is a P-adic integer. A necessary and sufficient condition that all elements of $\Omega(\zeta)$ with sufficiently large index belong to $M''(P)$, is that ζ is a rational number.*

PROOF: A) The condition is sufficient. Suppose that Ω_n belongs to $M''(P)$, say for all indices $n \geqq m+1$. Put

$$\Omega_{m,n} = \Omega_{m+1}\Omega_{m+2} \cdots \Omega_n,$$

so that

$$\mathbf{T}_n = \mathbf{T}_m\Omega_{m,n}.$$

Obviously

$$\Omega_{m,n} = \begin{pmatrix} P^a & Q \\ 0 & P^b \end{pmatrix},$$

where $a \geqq 0$, $b \geqq 0$ and Q are rational integers, and where $a + b = n - m$. Since both z_m and z_n lie in F, and

$$z_n = P^{b-a}z_m - P^{-a}Q,$$

necessarily

(35): $$\lim_{n\to\infty} b = +\infty.$$

But

$$\begin{pmatrix} p_n & p_n' \\ q_n & q_n' \end{pmatrix} = \begin{pmatrix} p_m & p_m' \\ q_m & q_m' \end{pmatrix}\begin{pmatrix} P^a & Q \\ 0 & P^b \end{pmatrix} = \begin{pmatrix} P^a p_m & Qp_m + P^b p_m' \\ P^a q_m & Qq_m + P^b q_m' \end{pmatrix},$$

and therefore by §2

$$P^a(p_m + q_m\zeta) \equiv 0 \pmod{P^n}, \quad \text{i.e.} \quad p_m + q_m\zeta \equiv 0 \pmod{P^{m+b}}.$$

Hence, from (35),

$$\zeta = P\text{-adic} \lim_{n\to\infty}\left(-\frac{p_m}{q_m}\right) = -\frac{p_m}{q_m},$$

as was to be proved.

B) The condition is necessary. Suppose that $\zeta = -p/q$, where $q \geqq 1$, and $(p, q) = 1$. Since ζ is a P-adic integer, $(P, q) = 1$.

To every index n, there is an integer A_n, such that

$$0 \leqq A_n \leqq P^n - 1, \qquad p + qA_n \equiv 0 \pmod{P^n},$$

and a second integer r_n, for which

$$p + qA_n = r_n P^n.$$

Hence

(36): $$(q, r_n) = 1,$$

since every common divisor of q and r_n is a divisor of p.

From (36), the equation

(37): $$r_n q_n' - r_n' q = 1$$

has integer solutions; if $q_n' = q_n^*$, $r_n' = r_n^*$ is one of them, then the general solution is given by

$$q_n' = q_n^* + kq, \qquad r_n' = r_n^* + kr_n,$$

where k is an arbitrary integer. We assume that this integer k is chosen such that the real part of the complex number

(38): $$z_n = -\frac{q_n'}{q} + \frac{P^n}{q(p - q\lambda)}$$

satisfies the inequalities

$$-\tfrac{1}{2} \leqq \Re(z_n) < \tfrac{1}{2}.$$

Obviously the imaginary part of the complex number $\dfrac{1}{q(p - q\lambda)}$ is positive.

Hence, if n is sufficiently large, z_n is a point in F.

Now, if as in §1,

$$Z_n = \frac{A_n + \lambda}{P^n} = \frac{-(p - q\lambda) + r_n P^n}{qP^n},$$

then it is easily verified that

(39): $$Z_n = \frac{r_n z_n + r_n'}{qz_n + q_n'}.$$

Hence z_n is the z_n of §1 for all sufficiently large n.

From (38), for two consecutive indices n and $n + 1$,

(40): $$z_n = \frac{z_{n+1} + \alpha_{n+1}'}{P}, \quad \text{where} \quad \alpha_{n+1}' = \frac{q_{n+1}' - Pq_n'}{q}.$$

Now $A_{n+1} = A_n + a_n P^n$ with an integer a_n, hence

$$p + qA_n = r_n P^n \quad \text{and} \quad p + q(A_n + a_n P^n) = r_{n+1} P^{n+1},$$

and therefore

$$r_n = Pr_{n+1} - qa_n.$$

Substituting this for r_n in (37) and subtracting the same equation with $n+1$ in place of n,

$$r_{n+1}(Pq'_n - q'_{n+1}) + q(r'_{n+1} - q'a_n - r'_n) = 0,$$

and finally from $(r_{n+1}, q) = 1$,

$$Pq'_n - q'_{n+1} \equiv 0 \pmod{q}.$$

Hence α'_{n+1} in (40) is an integer, and the matrix

$$\Omega_{n+1} = \begin{pmatrix} 1 & \alpha'_{n+1} \\ 0 & P \end{pmatrix}$$

of the transformation (40) is indeed an element of $M''(P)$, as was to be proved.

COROLLARY: *For a rational P-adic integer* ζ,

(41): $$\lim_{n\to\infty} y_n = +\infty.$$

PROOF: Immediate from (38).

18. A characteristic property of an irrational number

As we shall now prove, formula (41) is not true for irrational P-adic numbers. Put

(42): $$y(\zeta) = \liminf_{n\to\infty} y_n,$$

so that

$$\frac{\sqrt{3}}{2} \leqq y(\zeta) \leqq Y(\zeta).$$

THEOREM 12: *If* ζ *is an irrational P-adic integer, then for an infinity of indices* n

$$y_n \leqq \sqrt{P},$$

and therefore

(43): $$y(\zeta) \leqq \sqrt{P}.$$

PROOF: By Theorem 11, there are arbitrarily large indices $n+1$, for which $\Omega_{n+1} = \begin{pmatrix} \alpha_{n+1} & \alpha'_{n+1} \\ \beta_{n+1} & \beta'_{n+1} \end{pmatrix}$ does not belong to $M''(P)$, so that $\beta_{n+1} \not\equiv 0$. Hence

$$z_n = \frac{\alpha_{n+1} z_{n+1} + \alpha'_{n+1}}{\beta_{n+1} z_{n+1} + \beta'_{n+1}}$$

can be written as

$$(\beta_{n+1} z_n - \alpha_{n+1})(\beta_{n+1} z_{n+1} + \beta'_{n+1}) = -P.$$

Therefore

$$\min\left(\left|z_n - \frac{\alpha_{n+1}}{\beta_{n+1}}\right|, \left|z_{n+1} + \frac{\beta'_{n+1}}{\beta_{n+1}}\right|\right) \leqq \frac{\sqrt{P}}{|\beta_{n+1}|},$$

or

$$\min(y_n, y_{n+1}) \leqq \frac{\sqrt{P}}{|\beta_{n+1}|},$$

and the statement follows immediately, since $|\beta_{n+1}| \geqq 1$.

The question arises, whether the constant $\sqrt{P}$ in Theorem 12 is the best possible. This question is answered in the affirmative by the following

THEOREM 13: *To every $\epsilon > 0$, there is an irrational P-adic integer ζ, for which*

$$y(\zeta) \geqq \sqrt{P} - \epsilon.$$

PROOF: Let g be a large positive integer, and Ω the elliptic matrix

$$\Omega = \begin{pmatrix} 0 & P^{g+1} \\ -P^g & -1 \end{pmatrix}$$

of determinant P^{2g+1}; we suppose that g is so large that the fix point

$$f_1 = \frac{-1 + (1 - 4P^{2g+1})^{\frac{1}{2}}}{2P^g}$$

of Ω in F satisfies the inequality

$$|f_1 - \sqrt{-P}| \leqq \delta,$$

where δ is a positive constant to be assigned later.

By Theorem 3, there is a P-adic integer ζ_0 and a positive integer m, such that if $\mathrm{T}_m = \begin{pmatrix} p_m & p'_m \\ q_m & q'_m \end{pmatrix}$ and z_m are the elements of $\mathrm{T}(\zeta)$ and $z(\zeta)$ of index m, then

(44): $$q'_m \not\equiv 0 \pmod{P}, \qquad |z_m - f_1| \leqq \delta.$$

Put

$$\Omega^*_{m+k(2g+1)+h} = \left\{ \begin{array}{ll} \begin{pmatrix} 0 & P \\ -1 & 0 \end{pmatrix} & \text{for } h = 1 \\ \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} & \text{for } h = 2, 3, \cdots, g \\ \begin{pmatrix} 1 & 1 \\ 0 & P \end{pmatrix} & \text{for } h = g+1 \\ \begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix} & \text{for } h = g+2, g+3, \cdots, 2g+1 \end{array} \right\} \text{and } k = 0, 1, 2, \cdots,$$

and for $n \geqq m+1$

$$\mathrm{T}^*_n = \begin{pmatrix} p^*_n & p^{*\prime}_n \\ q^*_n & q^{*\prime}_n \end{pmatrix} = \mathrm{T}_m \Omega^*_{m+1} \Omega^*_{m+2} \cdots \Omega^*_n.$$

Since

$$\Omega^*_{m+k(2g+1)+1}\Omega^*_{m+k(2g+1)+2}\cdots\Omega^*_{m+(k+1)(2g+1)} = \begin{pmatrix} 0 & P \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix}^{g-1}\begin{pmatrix} 1 & 1 \\ 0 & P \end{pmatrix}\begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix}^{g} = \Omega,$$

obviously

$$\mathrm{T}^*_{m+k(2g+1)+h} = \mathrm{T}_m\Omega^k\Omega^*_{m+k(2g+1)+1}\cdots\Omega^*_{m+k(2g+1)+h},$$

and in particular

$$\mathrm{T}^*_{m+k(2g+1)} = \mathrm{T}_m\Omega^k.$$

Hence, by formula (26) in §9,

$$\mathrm{T}^*_{m+k(2g+1)} \equiv (-1)^{k-1}\mathrm{T}_m\Omega \equiv (-1)^k\begin{pmatrix} 0 & p'_m \\ 0 & q'_m \end{pmatrix} \pmod{P}.$$

From the first condition (44), therefore, $(q^*_n, q^{*\prime}_n) = 1$ for $n = m + k(2g + 1)$, and this must remain true also for all other indices $n \geqq m + 1$, since T^*_n is a left-hand divisor of all $\mathrm{T}^*_{m+k(2g+1)}$ with greater index.

Let z^*_n be the complex number

$$z^*_n = (\Omega^*_{m+1}\Omega^*_{m+2}\cdots\Omega^*_n)^{-1}z_m = \mathrm{T}^{*-1}_n\lambda \qquad (n \geqq m + 1),$$

so that in particular

$$z^*_{m+k(2g+1)} = \Omega^{-k}z_m \qquad \text{and} \qquad z^*_{m+k(2g+1)+1} = \frac{-P}{z^*_{m+k(2g+1)}}.$$

As in §10, since Ω has the fix point f_1,

$$|z^*_{m+k(2g+1)} - f_1| \leqq \delta_1 \qquad \text{for } k = 0, 1, 2, \cdots,$$

and therefore

$$\left|z^*_{m+k(2g+1)+1} + \frac{P}{f_1}\right| \leqq \delta_2 \qquad \text{for } k = 0, 1, 2, \cdots,$$

where both $\delta_1 > 0$ and $\delta_2 > 0$ only depend on δ and tend to zero with δ. In the second inequality

$$-\frac{P}{f_1} = \frac{1 + (1 - 4P^{2g+1})^{\frac{1}{2}}}{2P^g}.$$

Suppose now that $\epsilon \leqq \frac{1}{3}$, and that δ is so small that

$$\max(\delta + \delta_1, \delta + \delta_2) \leqq \epsilon.$$

Then, from the definition of g and from the last formulae,

$$|z^*_{m+k(2g+1)+h} - \sqrt{-P}| \leqq \epsilon \qquad \text{for } h = 0 \text{ or } 1,\ k = 0, 1, 2, \cdots,$$

so that in particular all points

$$z^*_{m+k(2g+1)}, \qquad z^*_{m+k(2g+1)+1}$$

lie in F, and their imaginary parts satisfy

(45): $\quad \Im(z^*_{m+k(2g+1)+h}) \geqq \sqrt{P} - \epsilon \qquad$ for $h = 0$ or 1, $k = 0, 1, 2, \cdots$.

The remaining z^*_n are given by

$$z^*_{m+k(2g+1)+h} = \left.\begin{cases} P^{h-1} z^*_{m+k(2g+1)+1} & \text{for } h = 2, 3, \cdots, g \\ P^g z^*_{m+k(2h+1)+1} - 1 & \text{for } h = g+1 \\ P^{2g-h+1} z^*_{m+k(2g+1)+1} - P^{g-h-1} & \\ \qquad \text{for } h = g+2, g+3, \cdots, 2g & \end{cases} \right\} \text{and } k = 0, 1, 2, \cdots,$$

so that for these numbers

(46): $\quad \Im(z^*_{m+k(2g+1)+h}) \geqq P(\sqrt{P} - \epsilon) \qquad$ for $h = 2, 3, \cdots, 2g$, $k = 0, 1, 2, \cdots$.

Suppose now that for $n \geqq m+1$, z_n is the point in F equivalent to z^*_n with respect to Γ, and that

$$z^*_n = \mathrm{P}_n z_n ,$$

where P_n belongs to Γ. Evidently

$$z_{m+k(2g+1)+h} = z^*_{m+k(2g+1)+h} , \qquad \mathrm{P}_{m+k(2g+1)+h} = \mathrm{E} \quad \text{for } h = 0 \text{ or } 1, k = 0, 1, 2, \cdots,$$

and

(47): $\qquad \Im(z_n) = \Im(z^*_n) \qquad$ for $n \geqq m+1$.

Put

$$\Omega_{m+1} = \Omega^*_{m+1} \mathrm{P}_{m+1} = \Omega^*_{m+1} , \qquad \Omega_n = \mathrm{P}^{-1}_{n-1} \Omega^*_n \mathrm{P}_n \qquad \text{for } n \geqq m+2,$$

and

$$\mathrm{T}_n = \begin{pmatrix} p_n & p'_n \\ q_n & q'_n \end{pmatrix} = \mathrm{T}_m \Omega_{m+1} \Omega_{m+2} \cdots \Omega_n \qquad \text{for } n \geqq m+1,$$

so that

$$\mathrm{T}_n = \mathrm{T}^*_n \mathrm{P}_n ,$$

and therefore in particular

$$\mathrm{T}_{m+k(2g+1)+h} = \mathrm{T}^*_{m+k(2g+1)+h} \qquad \text{for } h = 0 \text{ or } 1, k = 0, 1, 2, \cdots.$$

Since $(q^*_n , q^{*\prime}_n) = 1$ and T_n is a left-hand divisor of $\mathrm{T}^*_n = \mathrm{T}_n \mathrm{P}^{-1}_n$, obviously also

$$(q_n , q'_n) = 1 \qquad (n \geqq m+1).$$

Furthermore, all numbers

$$z_n = \mathrm{T}^{-1}_n \lambda$$

lie in F, as follows from the construction of z_n and T_n.

Hence, if

$$\{\Omega_1, \Omega_2, \cdots, \Omega_m\}, \qquad \{\mathrm{T}_0, \mathrm{T}_1, \cdots, \mathrm{T}_m\}, \qquad \{z_0, z_1, \cdots, z_m\}$$

are the elements of $\Omega(\zeta_0)$, $\mathrm{T}(\zeta_0)$, and $z(\zeta_0)$ with indices $n \leqq m$, we can apply Theorem 2 to the infinite sequences

$$\{\Omega_1, \Omega_2, \Omega_3, \cdots\}, \qquad \{\mathrm{T}_0, \mathrm{T}_1, \mathrm{T}_2, \cdots\}, \qquad \{z_0, z_1, z_2, \cdots\}.$$

We obtain the existence of a P-adic integer ζ (for which, by the way, $|\zeta - \zeta_0|_P \leqq P^{-m}$), such that $\{\Omega_n\} = \Omega(\zeta)$, $\{\mathrm{T}_n\} = \mathrm{T}(\zeta)$, $\{z_n\} = z(\zeta)$. This number is irrational by Theorem 12, since

$$\Omega_{m+k(2g+1)+1} = \Omega^*_{m+k(2g+1)+1} = \begin{pmatrix} 0 & P \\ -1 & 0 \end{pmatrix} \qquad (k = 0, 1, 2, \cdots)$$

does not belong to $M''(P)$. (Formula (27) in §9 shows that ζ lies in a quadratic field; compare the proof of the next theorem.) Also from (45), (46), and (47),

$$y_n \geqq \sqrt{P} - \epsilon,$$

for all $n \geqq m$, so that ζ satisfies all conditions of Theorem 13.

19. Additions to the last theorems

THEOREM 14: *If ζ is the P-adic integer of Theorem 13, then*

$$\sqrt{P} - \epsilon \leqq y_n \leqq \sqrt{P}, \qquad |p_n + q_n\zeta|_P = P^{-n}$$

for an infinity of indices.

PROOF: We specialize the formulae in §9 by taking

$$\Omega = \begin{pmatrix} 0 & P^{g+1} \\ -P^g & -1 \end{pmatrix}, \qquad \mathrm{T} = \mathrm{T}_m = \begin{pmatrix} p_m & p'_m \\ q_m & q'_m \end{pmatrix},$$

$$\mathrm{T}^* = \mathrm{T}_{m+k(2g+1)} = \begin{pmatrix} p_{m+k(2g+1)} & p'_{m+k(2g+1)} \\ q_{m+k(2g+1)} & q'_{m+k(2g+1)} \end{pmatrix},$$

where these matrices are the same as those defined in the proof of the last theorem. The equation for φ_1 and φ_2 becomes

$$\varphi^2 + \varphi + P^{2g+1} = 0,$$

and its two roots satisfy

$$|\varphi_1|_P = 1, \qquad |\varphi_2|_P = P^{-(2g+1)}.$$

By (27),

$$p_{m+k(2g+1)} - \frac{p_m P^{-g}\varphi_2 + p'_m}{q_m P^{-g}\varphi_2 + q'_m}\, q_{m+k(2g+1)} = \frac{\varphi_2^k P^m}{q_m P^{-g}\varphi_2 + q'_m}.$$

Here $|P^{-g}\varphi_2|_P \leqq 1/P$, and $|q'_m|_P = 1$ by construction; hence

$$\zeta = -\frac{p_m P^{-g}\varphi_2 + p'_m}{q_m P^{-g}\varphi_2 + q'_m}$$

is a P-adic integer, for which

$$(48): \quad |\, p_{m+k(2g+1)} + q_{m+k(2g+1)}\,\zeta\,|_P = \left| \frac{\varphi_2^k P^m}{q_m P^{-g}\varphi_2 + q_m'} \right|_P = P^{-\{m+k(2g+1)\}};$$

i.e. ζ is the number of Theorem 13.

All points $z_{m+k(2g+1)}$ lie on a circle K, whose points z satisfy the inequalities

$$|\, z - \sqrt{-P}\,| \leqq \epsilon;$$

since f_1 is an inner point of K and since $\Im(f_1) < \sqrt{P}$, K must have an arc C, for whose points $\Im(z) \leqq \sqrt{P}$. Now the multiplicator ϑ of Ω is not a root of unity, if, as we assume, g is sufficiently large. Hence there are arbitrarily large integer values of k, for which $z_{m+k(2g+1)}$ lies on C, so that the theorem follows immediately.

The number ζ of the last two theorems had the property that all elements of $\Omega(\zeta)$ with sufficiently large indices were either equal to $\mp\begin{pmatrix} 0 & P \\ -1 & 0 \end{pmatrix}$ or belonged to $M''(P)$. This is not accidental, as the following theorem shows:

THEOREM 15: *If an infinity of elements of $\Omega(\zeta)$ belong to $M'(P)$, but are different from $\mp\begin{pmatrix} 0 & P \\ -1 & 0 \end{pmatrix}$, then*

$$y(\zeta) \leqq \frac{(4P-1)^{\frac{1}{2}}}{2} < \sqrt{P}.$$

To every $\epsilon > 0$, there is a number ζ with this property, for which

$$\frac{(4P-1)^{\frac{1}{2}}}{2} - \epsilon \leqq y(\zeta) \leqq \frac{(4P-1)^{\frac{1}{2}}}{2}.$$

PROOF: Let $\Omega_{n+1} = \begin{pmatrix} \alpha_{n+1} & \alpha'_{n+1} \\ \beta_{n+1} & \beta'_{n+1} \end{pmatrix} \neq \mp\begin{pmatrix} 0 & P \\ -1 & 0 \end{pmatrix}$ belong to $M'(P)$, so that $\beta_{n+1} \neq 0$ and

$$(\beta_{n+1}z_n - \alpha_{n+1})(\beta_{n+1}z_{n+1} + \beta'_{n+1}) = -P.$$

If $|\, \beta_{n+1}\,| \geqq 2$, then, as in §18,

$$\min(y_n, y_{n+1}) \leqq \frac{\sqrt{P}}{|\beta_{n+1}|} < \frac{(4P-1)^{\frac{1}{2}}}{2}.$$

If $|\, \beta_{n+1}\,| = 1$, $\alpha_{n+1}\beta'_{n+1} \neq 0$, then $|\, \beta_{n+1}x_n - \alpha_{n+1}\,| \geqq \frac{1}{2}$, $|\, \beta_{n+1}x_{n+1} - \beta'_{n+1}\,| \geqq \frac{1}{2}$; since $\min(|\, z_n\,|, |\, z_{n+1}\,|) \leqq \sqrt{P}$,

$$\min(y_n, y_{n+1}) \leqq ((\sqrt{P})^2 - (\tfrac{1}{2})^2)^{\frac{1}{2}} = \frac{(4P-1)^{\frac{1}{2}}}{2}.$$

Finally, if $|\, \beta_{n+1}\,| = 1$ and $\alpha_{n+1}\beta'_{n+1} = 0$, then Ω_{n+1} is one of the matrices

$$\mp\begin{pmatrix} 0 & 1 \\ -P & a \end{pmatrix}, \quad \mp\begin{pmatrix} -a & 1 \\ -P & 0 \end{pmatrix}, \quad \mp\begin{pmatrix} 0 & P \\ -1 & a \end{pmatrix}, \quad \mp\begin{pmatrix} -a & P \\ -1 & 0 \end{pmatrix},$$

where a is an integer. Of these matrices, $\mp\begin{pmatrix} 0 & 1 \\ -P & a \end{pmatrix}$ and $\mp\begin{pmatrix} -a & 1 \\ -P & 0 \end{pmatrix}$ do not belong to $M(P)$. If $\Omega_{n+1} = \mp\begin{pmatrix} 0 & P \\ -1 & a \end{pmatrix}$, then $a \neq 0$ and $z_n = -P/(z_{n+1} - a)$; hence

$$y_n = \frac{Py_{n+1}}{(x_{n+1} - a)^2 + y_{n+1}^2} \leqq \frac{4Py_{n+1}}{1 + 4y_{n+1}^2}$$

since $| x_{n+1} - a | \geqq \frac{1}{2}$, and therefore

$$y_n - \frac{(4P-1)^{\frac{1}{2}}}{2} \leqq \frac{-2(4P-1)^{\frac{1}{2}}}{1 + 4y_{n+1}^2}\left(y_{n+1} - \frac{(4P-1)^{\frac{1}{2}}}{2}\right)\left(y_{n+1} - \frac{1}{2(4P-1)^{\frac{1}{2}}}\right),$$

so that again $\min(y_n, y_{n+1}) \leqq \frac{(4P-1)^{\frac{1}{2}}}{2}$. A similar proof applies to $\Omega_{n+1} = \mp\begin{pmatrix} -a & 1 \\ -P & 0 \end{pmatrix}$.

The second part of the Theorem is obtained immediately from Theorem 5, by taking $\Omega = \begin{pmatrix} 0 & P \\ -1 & 1 \end{pmatrix}$.

20. An application of the preceding results

For complex $z = x + yi$, let

$$t(z) = \frac{(x^2 + y^2)((1 - |x|)^2 + y^2)}{4y^3},$$

and, if ζ is a P-adic integer, define

$$T(\zeta) = \liminf_{n \to \infty} t(z_n).$$

The results in §§17–19 lead to the following

THEOREM 16: *If ζ is rational, then*

$$T(\zeta) = \infty.$$

If ζ is irrational, then for an infinity of indices

$$t(z_n) \leqq \frac{P+1}{4\sqrt{P}},$$

and therefore

$$T(\zeta) \leqq \frac{P+1}{4\sqrt{P}}.$$

Finally, to every $\epsilon > 0$, there is a P-adic integer ζ (which evidently must be irrational), for which

$$\frac{P+1}{4\sqrt{P}} - \epsilon \leqq T(\zeta) \leqq \frac{P+1}{4\sqrt{P}}.$$

PROOF: The first part of the theorem follows immediately from the corollary to Theorem 11, since x_n is bounded.

Now let ζ be irrational. The identity

$$(x^2+y^2)((1-|x|)^2+y^2)-y^2(1+y^2)=|x|(1-|x|)\{|x|(1-|x|)-2y^2\}$$

shows that for all z in F

$$(x^2+y^2)((1-|x|)^2+y^2) \leqq y^2(1+y^2),$$

and therefore

$$t(z) \leqq \frac{1+y^2}{4y}.$$

The function $(1+y^2)/4y$ has its minimum at $y=1$, and increases with increasing $|y-1|$. Hence, by Theorem 12, for an infinity of n

$$t(z_n) \leqq \max\left(\frac{1+\left(\frac{\sqrt{3}}{2}\right)^2}{4\left(\frac{\sqrt{3}}{2}\right)}, \frac{1+(\sqrt{P})^2}{4\sqrt{P}}\right)=\frac{P+1}{4\sqrt{P}},$$

since $(P+1)/4\sqrt{P}$ increases with P, and

$$\frac{7}{8\sqrt{3}} \leqq \frac{3}{4\sqrt{2}}.$$

The last part of the theorem follows from the proof of Theorem 13, by which, to every $\delta>0$ there is a P-adic integer ζ such that

$$|z_n-\sqrt{-P}| \leqq \delta$$

for an infinity of n, and

$$y_n \geqq P(\sqrt{P}-\delta)$$

for all other sufficiently large indices. For, if δ is chosen sufficiently small, then for the indices of the first class

$$\left|t(z_n)-\frac{P+1}{4\sqrt{P}}\right| \leqq \epsilon,$$

and for those of the second class

$$t(z_n) \geqq \frac{y_n}{4} \geqq \frac{P^{3/2}}{4}-\epsilon > \frac{P+1}{4\sqrt{P}}-\epsilon,$$

as was to be proved.—

For $P \leqq 19$, the expression $(P+1)/4\sqrt{P}$ is given in the following table:

P	2	3	5	7	11	13	17	19
$(P+1)/4\sqrt{P}$	$\frac{3}{4\sqrt{2}}$	$\frac{1}{\sqrt{3}}$	$\frac{3}{2\sqrt{5}}$	$\frac{2}{\sqrt{7}}$	$\frac{3}{\sqrt{11}}$	$\frac{7}{2\sqrt{13}}$	$\frac{9}{2\sqrt{17}}$	$\frac{5}{\sqrt{19}}$

If ζ satisfies the conditions of Theorem 15, then the upper bound for $T(\zeta)$ can be improved to

$$T(\zeta) \leqq \frac{4P+3}{8(4P-1)^{\frac{1}{2}}},$$

but this is probably not the best possible result.

FINAL REMARK: It is clear that if $f(x, y)$ is an arbitrary function, then the limits

$$\liminf_{n\to\infty} f(x_n, y_n), \qquad \limsup_{n\to\infty} f(x_n, y_n)$$

can be investigated in the same way as for the special cases we have considered. It would be useful to consider in particular the cases $f(x, y) = x/y$ and $f(x, y) = \frac{x^2+y^2}{y}$; the formulae (52) in the next paragraph show why these functions are of interest.

PART II: DIOPHANTINE APPROXIMATIONS IN THE FIELD OF THE P-ADIC NUMBERS

21. Introduction of a binary quadratic form

We shall now use the results of the first part to obtain theorems about *Diophantine Approximations* in the P-adic field. This problem is connected in a very simple way with the properties of the representative $z(\zeta)$ by means of the theory of positive definite binary and ternary quadratic forms.

As usual, if $z = x + yi$ is an arbitrary complex number, we denote its complex conjugate number by a bar: $\bar{z} = x - yi$. Let λ be the same number as in the first part, and $\Phi(X, Y)$ the positive definite binary quadratic form

$$\Phi(X, Y) = \frac{2(X-\lambda Y)(X-\bar{\lambda}Y)}{|\lambda-\bar{\lambda}|} = \mathbf{A}X^2 + 2\mathbf{B}XY + \mathbf{C}Y^2.$$

Obviously its determinant

$$\mathbf{B}^2 - \mathbf{AC} = \frac{(\lambda-\bar{\lambda})^2}{|\lambda-\bar{\lambda}|^2} = -1,$$

and since λ lies in F, Φ is a reduced form. In particular

$$\Phi(X, Y) = X^2 + Y^2, \qquad \text{if } \lambda = i,$$

$$\Phi(X, Y) = \frac{2(X^2+XY+Y^2)}{\sqrt{3}}, \qquad \text{if } \lambda = \frac{-1+\sqrt{-3}}{2}.$$

For $n = 0, 1, 2, \cdots$ put

(49): $$\Phi_n(X, Y) = \Phi(p_nX + p_n'Y, q_nX + q_n'Y) = \mathbf{A}_nX^2 + 2\mathbf{B}_nXY + \mathbf{C}_nY^2,$$

so that the determinant of Φ_n is

(50): $$\mathbf{B}_n^2 - \mathbf{A}_n\mathbf{C}_n = -P^{2n}.$$

By formula (8) in §2,

$$\Phi_n(X, Y) = \mathbf{A}_n(X - z_nY)(X - \bar{z}_nY), \tag{51}$$

and therefore from (50), if $z_n = x_n + y_n i$,

$$\mathbf{A}_n = \frac{P^n}{y_n}, \qquad \mathbf{B}_n = -\frac{P^n x_n}{y_n}, \qquad \mathbf{C}_n = \frac{P^n(x_n^2 + y_n^2)}{y_n}. \tag{52}$$

On the other hand, from (49):

$$\begin{aligned} \mathbf{A}_n &= \Phi(p_n, q_n), \\ \mathbf{B}_n &= \mathbf{A}p_np_n' + \mathbf{B}(p_nq_n' + p_n'q_n) + \mathbf{C}q_nq_n', \\ \mathbf{C}_n &= \Phi(p_n', q_n'). \end{aligned} \tag{53}$$

By construction, z_n lies in F. Hence all forms

$$\Phi_n(X, Y) \qquad (n = 0, 1, 2, \cdots)$$

are reduced, so that

$$-\mathbf{A}_n \leqq 2\mathbf{B}_n < \mathbf{A}_n \leqq \mathbf{C}_n, \quad P^{2n} \leqq \mathbf{A}_n\mathbf{C}_n \leqq \tfrac{4}{3}P^{2n}, \quad \mathbf{A}_n \leqq \frac{2}{\sqrt{3}}P^n. \tag{54}$$

Therefore, in particular:

THEOREM 17: *For any P-adic integer ζ and for every positive integer n, there are two integers p_n and q_n, such that*

$$|\, p_n + q_n\zeta \,|_P \leqq P^{-n}, \qquad 0 < \Phi(p_n, q_n) \leqq \frac{2}{\sqrt{3}}P^n.$$

22. The P-adic analogue to the Theorem of Lagrange

The following considerations depend on the identity

$$\Phi(p_nX + p_n'Y, q_nX + q_n'Y) = \frac{P^n}{y_n}\{X^2 - 2x_nXY + (x_n^2 + y_n^2)Y^2\}, \tag{55}$$

which follows immediately from (52) and (53), and on the inequality

$$|\, XY' - X'Y \,|^2 \leqq \Phi(X, Y)\Phi(X', Y'), \tag{56}$$

which is obvious from the identity

$$\begin{aligned} (\mathbf{AC} - \mathbf{B}^2)(XY' - X'Y)^2 \\ = \Phi(X, Y)\Phi(X', Y') - \{\mathbf{A}XX' + \mathbf{B}(XY' + X'Y) + \mathbf{C}YY'\}^2. \end{aligned}$$

THEOREM 18: *Let ζ be a P-adic integer, $\mathrm{T}_n = \begin{pmatrix} p_n & p_n' \\ q_n & q_n' \end{pmatrix}$ the element of $\mathrm{T}(\zeta)$ with index n, and p, q two integers, such that*

$$|\, p + q\zeta \,|_P \leqq P^{-n}, \qquad \Phi(p, q) > 0.$$

Then

$$\Phi(p, q) \geqq \Phi(p_n, q_n),$$

with equality if and only if either

$$|z_n| > 1 \quad \textit{and} \quad p = \eta p_n, \qquad q = \eta q_n, \qquad (\eta = \mp 1),$$

or

$$\left.\begin{aligned} |z_n| = 1, \qquad y_n > \frac{\sqrt{3}}{2} \quad \textit{and} \quad p = \eta p_n, \qquad q = \eta q_n, \quad \textit{or} \\ p = \eta p_n', \qquad q = \eta q_n', \end{aligned}\right\} (\eta = \mp 1),$$

or

$$\left.\begin{aligned} z_n = \frac{-1 + \sqrt{-3}}{2} \quad \textit{and} \quad p = \eta p_n, \qquad q = \eta q_n, \quad \textit{or} \\ p = \eta p_n', \quad q = \eta q_n', \quad \textit{or} \quad p = \eta(p_n - p_n'), \quad q = \eta(q_n - q_n'), \end{aligned}\right\} (\eta = \mp 1).$$

PROOF: Suppose that

$$|p + q\zeta|_P \leqq P^{-n}, \qquad 0 < \Phi(p, q) \leqq \Phi(p_n, q_n).$$

Then, from (56),

$$|pq_n' - p_n'q|^2 \leqq \Phi(p_n, q_n)\Phi(p_n', q_n') \leqq \tfrac{4}{3} P^{2n},$$

$$|p_nq - pq_n|^2 \leqq \Phi(p_n, q_n)\Phi(p_n, q_n) \leqq \tfrac{4}{3} P^{2n},$$

while on the other hand

$$|pq_n' - p_n'q|_P$$

$$= |q_n'(p + q\zeta) - q(p_n' + q_n'\zeta)|_P \leqq \max(|p + q\zeta|_P, |p_n' + q_n'\zeta|_P) \leqq P^{-n},$$

$$|p_nq - pq_n|_P$$

$$= |q(p_n + q_n\zeta) - q_n(p + q\zeta)|_P \leqq \max(|p_n + q_n\zeta|_P, |p + q\zeta|_P) \leqq P^{-n}.$$

Therefore each of the two numbers

$$X = \frac{pq_n' - p_n'q}{p_nq_n' - p_n'q_n} \quad \text{and} \quad Y = \frac{p_nq - pq_n}{p_nq_n' - p_n'q_n}$$

is equal either to 0, or to $+1$, or to -1. Now obviously

$$p = p_nX + p_n'Y, \qquad q = q_nX + q_n'Y,$$

so that one of the following four cases occur, where $\eta = \mp 1$:

a) $p = \eta p_n$, $\quad q = \eta q_n$. $\quad$ b) $p = \eta p_n'$, $\quad q = \eta q_n'$.

c) $p = \eta(p_n + p_n')$, $\quad q = \eta(q_n + q_n')$. $\quad$ d) $p = \eta(p_n - p_n')$, $\quad q = \eta(q_n - q_n')$.

The first case is trivial. In case b), from (52),

$$\Phi(p, q) = \Phi(p_n', q_n') = \frac{P^n(x_n^2 + y_n^2)}{y_n} \leqq \Phi(p_n, q_n) = \frac{P^n}{y_n},$$

so that necessarily $| z_n | = 1$. In case c), from (55),

$$\Phi(p, q) = \Phi(p_n + p_n', q_n + q_n') = \frac{P^n(1 - 2x_n + x_n^2 + y_n^2)}{y_n} \leqq \Phi(p_n, q_n) = \frac{P^n}{y_n},$$

which is impossible for all points z_n in F. Finally in case d), from (55),

$$\Phi(p, q) = \Phi(p_n - p_n', q_n - q_n') = \frac{P^n(1 + 2x_n + x_n^2 + y_n^2)}{y_n} \leqq \Phi(p_n, q_n) = \frac{P_n}{y_n},$$

which requires that $z_n = \dfrac{-1 + \sqrt{-3}}{2}$.

23. P-adic analogues to the Theorem of Hurwitz-Borel

The results in §§12–16, the formulae in the last two paragraphs, and Theorem 18 give immediately the following theorems, which form improvements on Theorem 17:

THEOREM 19: A) *To every diadic integer ζ and for at least one of any three consecutive indices n, there are two integers p_n, q_n, for which*

$$| p_n + q_n\zeta |_2 \leqq 2^{-n}, \qquad 0 < \Phi(p_n, q_n) \leqq \frac{2}{\sqrt{7}} \cdot 2^n.$$

B) *To every $\epsilon > 0$, there is a diadic integer ζ, such that*

$$| p + q\zeta |_2 \leqq 2^{-n}, \qquad 0 < \Phi(p, q) \leqq \left(\frac{2}{\sqrt{7}} - \epsilon\right) \cdot 2^n$$

has no integer solution p, q for sufficiently large n.

THEOREM 20: A) *To every triadic integer ζ and for at least one of any three consecutive indices n, there are two integers p_n, q_n, for which*

$$| p_n + q_n\zeta |_3 \leqq 3^{-n}, \qquad 0 < \Phi(p_n, q_n) \leqq \frac{1}{\sqrt{2}} \cdot 3^n.$$

B) *To every $\epsilon > 0$, there is a triadic integer ζ, such that*

$$| p + q\zeta |_3 \leqq 3^{-n}, \qquad 0 < \Phi(p, q) \leqq \left(\frac{1}{\sqrt{2}} - \epsilon\right) \cdot 3^n$$

has no integer solution p, q for sufficiently large n.

THEOREM 21: A) *To every pentadic integer ζ and for at least one of any two consecutive indices n, there are two integers p_n, q_n, for which*

$$| p_n + q_n\zeta |_5 \leqq 5^{-n}, \qquad 0 < \Phi(p_n, q_n) \leqq 5^n.$$

B) *To every* $\epsilon > 0$, *there is a pentadic integer* ζ, *such that*

$$|\, p + q\zeta \,|_5 \leqq 5^{-n}, \qquad 0 < \Phi(p, q) \leqq (1 - \epsilon)\cdot 5^n$$

has no integer solution p, q *for sufficiently large* n.

THEOREM 22: *If* $P \equiv 1 \pmod 6$, *then to every* $\epsilon > 0$ *there is a* P*-adic integer* ζ, *such that*

$$|\, p + q\zeta \,|_P \leqq P^{-n}, \qquad 0 < \Phi(p, q) \leqq \left(\frac{2}{\sqrt{3}} - \epsilon\right)\cdot P^n$$

has no integer solution p, q *for sufficiently large* n.

Probably for any prime number P and for any P-adic integer ζ

$$\liminf_{n\to\infty} \min_{\substack{|p+q\zeta|_P \leqq P^{-n} \\ \Phi(p,q)>0}} \frac{\Phi(p, q)}{P^n} \leqq \frac{1}{Y(P)},$$

where $Y(P)$ is the arithmetical function defined in §11; but so far I could prove this only for $P \leqq 7$. That $1/Y(P)$ cannot, in general, be replaced by any smaller number, follows from Theorem 7.

24. The P-adic analogue to the Theorem of Khintchine

The results in §§18–19, the formulae in §21, and Theorem 18 lead immediately to the following results:

THEOREM 23: A) *To every irrational* P*-adic integer* ζ *there is an infinity of indices* n *for which the conditions*

$$|\, p + q\zeta \,|_P \leqq P^{-n}, \qquad 0 < \Phi(p, q) < \frac{P^n}{\sqrt{P}}$$

have no integer solution p, q.

B) *To every* $\epsilon > 0$, *there is an irrational* P*-adic integer* ζ, *such that for all sufficiently large* n *there are two integers* p_n, q_n, *for which*

$$|\, p_n + q_n\zeta \,|_P \leqq P^{-n}, \qquad 0 < \Phi(p_n, q_n) \leqq \left(\frac{1}{\sqrt{P}} + \epsilon\right)P^n,$$

while for an infinity of indices

$$|\, p_n + q_n\zeta \,|_P = P^{-n}, \qquad \frac{P^n}{\sqrt{P}} \leqq \Phi(p_n, q_n) \leqq \left(\frac{1}{\sqrt{P}} + \epsilon\right)P^n.$$

THEOREM 24: A) *If an infinity of elements of* $\Omega(\zeta)$ *belong to* $M'(P)$, *but are different from* $\mp\begin{pmatrix} 0 & P \\ -1 & 0 \end{pmatrix}$, *then there is an infinity of indices* n *for which the inequalities*

$$|\, p + q\zeta \,|_P \leqq P^{-n}, \qquad 0 < \Phi(p, q) < \frac{2P^n}{(4P - 1)^{\frac{1}{2}}}$$

have no integer solution p, q.—

B) *To every* $\epsilon > 0$, *there is a P-adic integer* ζ *with the last property, such that for all sufficiently large* n *there are two integers* p_n, q_n, *for which*

$$|p_n + q_n\zeta|_P \leqq P^{-n}, \qquad 0 < \Phi(p_n, q_n) \leqq \left(\frac{2}{(4P-1)^{\frac{1}{4}}} + \epsilon\right) \cdot P^n.$$

25. The P-adic analogue to the Theorem of Tchebycheff

Ch. Hermite proved the following theorem:[10]

THEOREM 25: *Let*

$$\varphi(X, Y) = AX^2 + 2BXY + CY^2$$

be a reduced positive definite quadratic form, X_0 *and* Y_0 *two real numbers. Then there are two integers* X *and* Y, *such that*

$$\varphi(X - X_0, Y - Y_0) \leqq \frac{AC(A - 2|B| + C)}{4(AC - B^2)}.$$

By means of this lemma and the results in §20, we shall prove:

THEOREM 26: *Let* ζ *and* ϑ *be two P-adic integers, of which* ζ *is irrational. Then there is an infinity of indices* n, *for which there are two integers* u_n *and* v_n, *such that*

$$|u_n + v_n\zeta + \vartheta|_P \leqq P^{-n}, \qquad \Phi(u_n, v_n) \leqq \frac{P+1}{4\sqrt{P}} \cdot P^n.$$

PROOF: For every index n let E_n be the integer defined by

$$|\vartheta - E_n|_P \leqq P^{-n}, \qquad 0 \leqq E_n \leqq P^n - 1.$$

By Theorem 25, there are two integers X_n and Y_n, such that

$$\Phi(p_nX_n + p_n'Y_n - E_n, q_nX_n + q_n'Y_n) \leqq \frac{\mathbf{A}_n\mathbf{C}_n(\mathbf{A}_n - 2|\mathbf{B}_n| + \mathbf{C}_n)}{4(\mathbf{A}_n\mathbf{C}_n - \mathbf{B}_n^2)},$$

since

$$\Phi(p_nX + p_n'Y, q_nX + q_n'Y) = \Phi_n(X, Y) = \mathbf{A}_nX^2 + 2\mathbf{B}_nXY + \mathbf{C}_nY^2$$

is a reduced positive definite form. Obviously, from (52),

$$\frac{\mathbf{A}_n\mathbf{C}_n(\mathbf{A}_n - 2|\mathbf{B}_n| + \mathbf{C}_n)}{4(\mathbf{A}_n\mathbf{C}_n - \mathbf{B}_n^2)} = P^n\,\frac{(x_n^2 + y_n^2)((1 - |x_n|)^2 + y_n^2)}{4y_n^3} = P^n t(z_n);$$

hence, by Theorem 16, for an infinite sequence of indices n

$$\Phi(p_nX_n + p_n'Y_n - E_n, q_nX_n + q_n'Y_n) \leqq \frac{P+1}{4\sqrt{P}} \cdot P^n.$$

[10] Hermite, l. c. 6, 94–99.

Now put

$$v_n = q_n X_n + q_n' Y_n, \qquad w_n = r_n X_n + r_n' Y_n, \qquad u_n = P^n w_n - A_n v_n - E_n,$$

where A_n, r_n, r_n' are defined as in §§1–2. Since

$$p_n = P^n r_n - A_n q_n, \qquad p_n' = P^n r_n' - A_n q_n',$$

obviously

$$u_n = P^n(r_n X_n + r_n' Y_n) - A_n(q_n X_n + q_n' Y_n) - E_n = p_n X_n + p_n' Y_n - E_n,$$

and therefore

$$\Phi(u_n, v_n) \leqq \frac{P+1}{4\sqrt{P}} P^n.$$

On the other hand

$$u_n + A_n v_n + E_n = P^n w_n, \quad \text{i.e.} \quad |u_n + \zeta v_n + \vartheta|_P \leqq P^{-n},$$

so that u_n and v_n have the required properties.—

If $\Omega(\zeta)$ contains an infinity of elements of $M'(P)$ different from $\mp\begin{pmatrix} 0 & P \\ -1 & 0 \end{pmatrix}$, then the constant $(P+1)/4\sqrt{P}$ in the last theorem can be replaced by the smaller number $(4P+3)/4(4P-1)^{\frac{1}{2}}$. It is nearly certain that both these constants are not the best possible ones; so far, however, I have not been able to obtain the true values.

University of Manchester,
Manchester, England.

ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

TABLE OF REDUCED POSITIVE QUATERNARY QUADRATIC FORMS

BY S. B. TOWNES

(Received March 7, 1939)

Charve[1] published a table of reduced positive quaternary quadratic forms, which, however, contains several pairs of equivalent forms. In 1936 the author found conditions for unique reduced forms according to the Eisenstein method.[2] The forms in the following table obey these conditions.

Forms in different genera are separated by semicolons, those in different orders by colons, properly and improperly primitive forms by periods. The symbol (*abcdrstlmn*) represents the form of matrix

$$\begin{pmatrix} a & t & s & l \\ t & b & r & m \\ s & r & c & n \\ l & m & n & d \end{pmatrix};$$

a sequence of equal coefficients is indicated by a subscript. The table contains all reduced forms of determinant $D \leqq 25$, with the exception of the only imprimitive form (2 2 2 2) of determinant 16.

D	
1	(1 1 1 1).
2	(1 1 1 2).
3	(1 1 1 3) ; (1 1 2 2 0_5 1).
4	(1 1 1 4) ; (1 2 2 2 1 0_3 1_2) ; (1 1 2 2). (2 2 2 2 1_5 0).
5	(1 1 1 5) ; (1 1 2 3 0_5 1) . (2 2 2 2 1_5).
6	(1 1 1 6) , (1 2 2 2 1 0_5) ; (1 1 2 3).
7	(1 1 1 7) , (1 1 2 4 0_5 1) ; (1 2 2 3 1 0_3 1_2).
8	(1 1 1 8) ; (1 1 3 3 0_5 1) ; (1 2 2 3 0_4 1_2) ; (2 2 2 3 1_5 0) : (1 1 2 4) : (1 2 2 2). (2 2 2 2 1_3 0_3) .
9	(1 1 1 9) , (2 2 2 3 1_6) ; (1 1 2 5 0_5 1) : (1 1 3 3) ; (1 2 2 3 1 0_5). (2 2 2 2 0_2 1 0_2 1).
10	(1 1 1 10) , (1 2 2 3 0_4 1 0) ; (1 1 2 5) , (1 2 2 4 1 0_3 1_2).
11	(1 1 1 11), (1 1 3 4 0_5 1); (1 1 2 6 0_5 1), (2 2 2 3 0_2 1_4).
12	(1 1 1 12) , (2 2 2 3 1_3 0_3) ; (1 1 3 4) , (1 2 2 4 0_4 1_2) ; (1 2 3 3 1 0_3 1_2) ; (1 2 2 4 1 0_5) : (1 1 2 6) ; (1 2 2 3) : (1 1 4 4 0_5 2) ; (2 2 2 3 0_3 1_3) . (2 2 2 2 0_2 1 0_3) ; (2 2 2 4 1_5 0).
13	(1 1 1 13) , (1 2 2 5 1 0_3 1_2) ; (1 1 2 7 0_5 1) , (1 2 3 3 1 0_4 1) . (2 2 2 4 1_6).

[1] Comptes Rendus (Paris), vol. 96 (1883), p. 773.

[2] Doctor's dissertation, Chicago, 1936 (unpublished).

D

14 (1 1 1 14) , (1 1 2 7) , (1 2 2 4 0_4 1 0) ; (1 1 3 5 0_5 1) , (2 2 2 3 0_2 1_3 0).

15 (1 1 1 15) , (1 1 4 4 0_5 1) , (2 2 2 3 0_2 1 0_2 1) ; (1 2 3 3 1 0_5);
(1 2 2 5 1 0_5) , (2 2 3 3 1_5 0) ; (1 1 3 5) , (1 1 2 8 0_5 1).

16 (1 1 1 16), (1 2 2 5 0_4 1_2) ; (1 2 2 6 1 0_3 1_2) ; (1 1 4 4 0_5 2),
(2 2 2 5 1_5 0) ; (2 2 3 3 1_6) : (1 1 2 8), (2 2 2 3 0_3 1_2 0) ; (1 2 3 3 0_5 1) :
(1 1 4 4), (2 2 3 3 1_2 0 1_3) : (1 2 2 4) : (1 3 3 3 -1 0_3 1_2) .
(2 2 2 4 1_3 0_3).

17 (1 1 1 17) , (1 1 2 9 0_5 1) , (2 2 2 5 1_6) ; (1 1 3 6 0_5 1) , (1 2 3 4 1 0_3 1_2).
(2 2 2 4 0_2 1_4).

18 (1 1 1 18) , (1 2 2 5 0_4 1 0) , (2 2 3 3 1_3 0_2 1) ; (1 1 2 9),
(1 2 3 4 1 0_4 1) : (1 1 3 6), (2 2 2 3 0_2 1 0_3) ; (1 2 3 3) , (1 2 2 6 1 0_5).

19 (1 1 1 19), (1 1 4 5 0_5 1) , (1 2 3 4 0_4 1_2) ; (1 1 2 10 0_5 1),
(1 2 2 7 1 0_3 1_2) , (2 2 3 3 1_2 0 1 0 1).

20 (1 1 1 20) , (1 1 4 5) ; (1 2 3 4 1 0_5) , (2 2 3 3 1_2 0_3 1) ; (1 2 2 6 0_4 1_2) ,
(2 2 2 5 1_3 0_3) ; (1 1 3 7 0_5 1) , (1 2 4 4 1 0_3 1 2) : (1 1 2 10) ,
(2 2 2 3 0_3 1 0_2) ; (1 2 2 5) , (1 1 4 6 0_5 2) : (1 3 3 3 1 0_3 1_2) .
(2 2 2 4 0_2 1_3 0) : (2 2 2 4 0_3 1_3) ; (2 2 2 6 1_5 0) .

21 (1 1 1 21) , (2 2 3 3 1_3 0_3) ; (1 1 3 7) , (1 2 3 4 0_4 1 0) ,
(2 2 3 3 0 1 0_2 1_2) ; (1 1 2 11 0_5 1) , (1 3 3 3 1 0_3 1 0) ; (1 1 5 5 0_5 2) ,
(1 2 2 7 1 0_5) . (2 2 4 0_2 1 0_2 1) ; (2 2 2 6 1_6) .

22 (1 1 1 22) , (1 2 2 6 0_4 1 0) , (1 2 2 8 1 0_3 1_2) , (1 2 3 5 1 0_3 1_2) ;
(1 1 2 11) , (1 2 3 4 0_5 1) , (2 2 3 4 1_5 0).

23 (1 1 1 23) , (1 1 2 12 0_5 1) , (1 1 3 8 0_5 1) , (1 1 4 6 0_5 1) ,
(2 2 2 5 0_2 1_4) ; (1 2 3 5 1 0_4 1) , (2 2 3 4 1_6) .

24 (1 1 1 24) , (2 2 3 3 0_2 1 0_2 1) ; (1 3 3 3 1 0_5) , (2 2 3 4 1_2 0 1_3) ;
(1 2 2 8 1 0_5) , (2 2 2 7 1_5 0) ; (1 1 3 8) , (1 3 3 4 -1 0_3 1_2) ;
(1 2 2 7 0_4 1_2) , (2 3 3 3 -1 -1 1 0 1_2) ; (1 1 5 5 0_5 1) ; (1 1 4 7 0_5 2) ,
(1 2 4 4 1 0_3 1_2) ; (2 2 3 3 1_2 0_4) : (1 1 2 12) , (1 2 3 4) ; (1 1 4 6) ,
(2 2 3 3 0 1 0 1 0 1) : (1 2 2 6) ; (2 2 2 3) , (1 2 4 4 0_5 2) .
(2 2 2 4 0_2 1 0_3) ; (2 2 2 4 0_3 1_2 0) : (2 2 2 6 1_3 0_3) .

25 (1 1 1 25) , (1 2 3 5 0_4 1_2) , (2 2 2 7 1_6) ; (1 1 2 13 0_5 1) ,
(1 2 2 9 1 0_3 1_2) , (2 2 3 4 1_3 0_2 1) : (1 1 5 5) , (2 2 3 3 0 1 0_2 1 0) ;
(1 2 3 5 1 0_5) . (2 2 4 4 -1_2 1_4) .

THE UNIVERSITY OF OKLAHOMA,
NORMAN, OKLAHOMA.

ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

NOTE ON LINEAR FRACTIONAL SUBSTITUTIONS WITH LARGE DETERMINANT

BY H. DAVENPORT

(Received December 30, 1938)

1

In the preceding paper, Dr. Mahler has defined an arithmetical function $Y(P)$ for any prime P. Let $\alpha, \beta, \alpha', \beta'$ be any integers satisfying $\alpha\beta' - \alpha'\beta = P$, such that one of the fixed points f of the substitution

$$Z = \frac{\alpha z + \alpha'}{\beta z + \beta'} \tag{1}$$

lies in the fundamental domain of the modular group. Then $Y(P)$ is defined by

$$Y(P) = \min (\Im f)$$

for all such substitutions. The object of this note is to prove that

$$Y(P) = \tfrac{1}{2}\sqrt{3} + O(P^{-(1/12)+\epsilon}) \tag{2}$$

as $P \to \infty$, for any $\epsilon > 0$. More generally, I prove the following

THEOREM. *Let z_0 be any complex number other than 0. If P is a large positive integer, there exists a substitution* (1) *of determinant P such that one of its fixed points f satisfies*

$$f = z_0 + O(P^{-(1/12)+\epsilon}). \tag{3}$$

Moreover, the signs of $\Re(f) - \Re(z_0)$ and of $|f| - |z_0|$ may be prescribed arbitrarily.

If we take $z_0 = \frac{1}{2} + \frac{1}{2}\sqrt{3}i$ in this theorem, and prescribe $\Re(f) - \Re(z_0)$ to be negative, and $|f| - |z_0|$ to be positive, then f lies in the fundamental domain of the modular group, and (3) implies (2).

The proof of the theorem is an exercise in the use of exponential sums.

2

LEMMA. *Let N be a large positive integer, and let P be an integer relatively prime to N. For any N_1, N_2 there exist integers u, v satisfying*

$$uv \equiv P \pmod{N}, \tag{4}$$

$$N_1 < u < N_1 + N^{(5/6)+\epsilon}, \qquad N_2 < v < N_2 + N^{(5/6)+\epsilon}. \tag{5}$$

PROOF. For brevity, we write $e(t) = e^{2\pi it/N}$. The expression

$$\frac{1}{N}\sum_{x=0}^{N-1}\sum_{N_1<y<N_1+N^{(5/6)+\epsilon}} e(x(u-y))$$

has the value 1 if u or a number congruent to u satisfies $N_1 < u < N_1 + N^{(5/6)+\epsilon}$, and 0 otherwise. Hence the number of solutions of the conditions (4), (5) is

$$\frac{1}{N^2}\sum_{\substack{u=1\\(u,N)=1}}^{N-1}\sum_{x_1=0}^{N-1}\sum_{x_2=0}^{N-1}\sum_{y_1}\sum_{y_2} e(x_1(u-y_1)+x_2(Pu^{-1}-y_2)),$$

where Pu^{-1} denotes any integer v satisfying $uv \equiv P \pmod N$, and y_1, y_2 are summed over integers satisfying

$$N_1 < y_1 < N_1 + N^{(5/6)+\epsilon}, \qquad N_2 < y_2 < N_2 + N^{(5/6)+\epsilon}.$$

The contribution of the terms with $x_1 = x_2 = 0$ is

$$\frac{\varphi(N)}{N^2}\sum_{y_1}\sum_{y_2} 1 \sim \varphi(N)N^{-\frac{1}{3}+2\epsilon}.$$

Let $S(a, b)$ denote the Kloosterman sum

$$S(a,b) = \sum_{\substack{u=1\\(u,N)=1}}^{N-1} e(au + bu^{-1}).$$

The contribution of the terms with $x_1 = 0$, $x_2 \neq 0$ is

$$\Sigma_1 = \frac{1}{N^2}\left(\sum_{y_1} 1\right)\left(\sum_{x_2=1}^{N-1} S(0, Px_2)\sum_{y_2} e(-x_2y_2)\right),$$

and the contribution Σ_1' of the terms with $x_1 \neq 0$, $x_2 = 0$ has a similar form. The contribution of the terms with $x_1 \neq 0$, $x_2 \neq 0$ is

$$\Sigma_2 = \frac{1}{N^2}\sum_{x_1=1}^{N-1}\sum_{x_2=1}^{N-1} S(x_1, Px_2)\sum_{y_1} e(-x_1y_1)\sum_{y_2} e(-x_2y_2).$$

It is known[1] that

$$S(a,b) = O(N^{\frac{1}{2}+\epsilon}(a, N)),$$

where on the right a may be replaced by b. Also, if $1 \leqq x \leqq N-1$,

$$\sum_{y} e(-xy) = O\left(\frac{N}{x'}\right)$$

[1] Salié, *Math. Zeitschrift*, 36 (1932), pp. 263–278. For the case of a prime N, from which the general case follows trivially, see also Davenport, *Journal für Math.*, 169 (1933), pp. 158–176.

where $x' = \min(x, N - x)$, and y is summed over any interval. Hence

$$\begin{aligned}\sum_1 &= O\left(N^{-2}N^{(5/6)+\epsilon}\sum_{x_2=1}^{N-1} N^{\frac{1}{2}+\epsilon}(Px_2, N)\frac{N}{x_2'}\right)\\ &= O\left(N^{\frac{1}{2}+2\epsilon}\sum_{x_2=1}^{N-1}(x_2, N)\frac{1}{x_2}\right)\\ &= O\left(N^{\frac{1}{2}+2\epsilon}\sum_{d|N} d\sum_{m<N/d}\frac{1}{md}\right)\\ &= O(N^{\frac{1}{2}+2\epsilon}\log N\, d(N)),\end{aligned}$$

where $d(N)$ denotes the number of divisors of N. Similarly, $\sum_1'$ is of the same order of magnitude. Finally,

$$\begin{aligned}\sum_2 &= O\left(N^{-2}\sum_{x_1=1}^{N-1}\sum_{x_2=1}^{N-1} N^{\frac{1}{2}+\epsilon}(x_1, N)\frac{N}{x_1'}\frac{N}{x_2'}\right)\\ &= O\left(N^{\frac{1}{2}+\epsilon}\log N\sum_{x_1=1}^{N-1}(x_1, N)\frac{1}{x_1'}\right)\\ &= O(N^{\frac{1}{2}+\epsilon}\log^2 N\, d(N)).\end{aligned}$$

The upper bounds for $\sum_1$, $\sum_1'$, $\sum_2$ are all of lower order of magnitude than $\varphi(N)N^{-\frac{1}{2}+2\epsilon}$, since

$$\log N = O(N^{\epsilon_1}), \qquad d(N) = O(N^{\epsilon_1}), \qquad \frac{N}{\varphi(N)} = O(N^{\epsilon_1})$$

for any $\epsilon_1 > 0$. Hence the number of solutions of (4), (5) is positive.

3. Proof of the Theorem

Suppose first that both the prescribed signs are positive. Let $r = \Re(z_0)$, $m = |z_0|^2$. Let k be any fixed number greater than m. Let

$$\lambda = r + (r^2 - m + k)^{\frac{1}{2}}, \qquad \mu = -r + (r^2 - m + k)^{\frac{1}{2}}, \tag{6}$$

where the radicals are taken positively. Clearly $\lambda > \mu > 0$.

Choose β to be one of the two integers next less than $(P/k)^{\frac{1}{2}}$, and to be relatively prime to P (at least one of any two consecutive integers is prime to P). We have

$$P \geqq k\beta^2, \qquad P = k\beta^2 + O(\beta). \tag{7}$$

By the Lemma, there exist integers α, β' such that

$$\lambda\beta - \beta^{(5/6)+\epsilon} < \alpha < \lambda\beta, \qquad \mu\beta - 2\beta^{(5/6)+\epsilon} < \beta' < \mu\beta - \beta^{(5/6)+\epsilon}, \tag{8}$$

$$\alpha\beta' \equiv P (\mathrm{mod}\ \beta). \tag{9}$$

By (9), there exists an integer α' such that

$$\alpha\beta' - \alpha'\beta = P. \tag{10}$$

By (8) and (6),

$$\alpha - \beta' > (\lambda - \mu)\beta = 2r\beta,$$
$$\alpha - \beta' = 2r\beta + O(\beta^{(5/6)+\epsilon}).$$

By (7), (8), (10), and (6),

$$\alpha\beta' < \lambda\mu\beta^2, \qquad \alpha\beta' = \lambda\mu\beta^2 + O(\beta^{(11/6)+\epsilon}),$$
$$P - \alpha\beta' > (k - \lambda\mu)\beta^2 = m\beta^2,$$
$$P - \alpha\beta' = m\beta^2 + O(\beta^{(11/6)+\epsilon}).$$

Hence

$$0 < \frac{\alpha - \beta'}{2\beta} - r = O(\beta^{-(1/6)+\epsilon}) = O(P^{-(1/12)+\epsilon}), \tag{11}$$

$$0 < \frac{-\alpha'}{\beta} - m = O(\beta^{-(1/6)+\epsilon}) = O(P^{-(1/12)+\epsilon}). \tag{12}$$

Since the fixed points of the substitution (1) are the roots of

$$\beta f^2 - (\alpha - \beta')f - \alpha' = 0,$$

we have

$$\Re f = \frac{\alpha - \beta'}{2\beta}, \qquad |f|^2 = -\frac{\alpha'}{\beta},$$

and the theorem follows from (11) and (12).

If the prescribed signs are $-$ for $\Re(f) - \Re(z_0)$ and $+$ for $|f| - |z_0|$, it is only necessary to replace the intervals (8) by

$$\lambda\beta - 2\beta^{(5/6)+\epsilon} < \alpha < \lambda\beta - \beta^{(5/6)+\epsilon}, \qquad \mu\beta - \beta^{(5/6)+\epsilon} < \beta' < \mu\beta.$$

If the prescribed sign for $|f| - |z_0|$ is negative, we choose β to be one of the two integers next greater than $(P/k)^{\frac{1}{2}}$, and take either the intervals

$$\lambda\beta + \beta^{(5/6)+\epsilon} < \alpha < \lambda\beta + 2\beta^{(5/6)+\epsilon}, \qquad \mu\beta < \beta' < \mu\beta + \beta^{(5/6)+\epsilon},$$

or the intervals

$$\lambda\beta < \alpha < \lambda\beta + \beta^{(5/6)+\epsilon}, \qquad \mu\beta + \beta^{(5/6)+\epsilon} < \beta' < \mu\beta + 2\beta^{(5/6)+\epsilon}.$$

THE UNIVERSITY, MANCHESTER, ENGLAND.

ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

BEWEIS EINER CARATHEODORYSCHEN VERMUTUNG. TEIL I

VON HANS HAMBURGER

(Received February 4, 1939)

1. Herr Caratheodory hat gelegentlich die Vermutung geäussert, dass eine überall reguläre geschlossene Fläche vom Geschlechte Null (mit stetiger Tangentenebene und stetiger Krümmung) mindestens zwei Nabelpunkte besitzt. Der Beweis dieser Vermutung wird unter gewissen weiter unten angegebenen Differenzierbarkeitsvoraussetzungen, die für analytische Flächen sicher alle erfüllt sind, in der vorliegenden Arbeit und einem später zu veröffentlichenden 2. Teile geführt werden.

Die Caratheodorysche Behauptung, die eine Aussage über die Differentialgeometrie im Grossen enthält, lässt sich nun aber in einfacher Weise auf eine Behauptung über das Verhalten der Krümmungslinien in der Umgebung eines isolierten Nabels reduzieren, wenn man den Begriff des topologischen Index eines singulären Punktes im Kurvennetz heranzieht.

2. Die topologischen Eigenschaften eines Kurvennetzes sind vom Verf. schon früher an anderer Stelle[1] untersucht worden. Ist auf einem ebenen Bereich oder einem Flächenstück ein Kurvennetz, das sind zwei sich schneidende Kurvenscharen, ausgebreitet, so heisst ein Punkt P ein regulärer Punkt des Netzes, wenn die Netzkurven, die eine hinreichend kleine Umgebung des Punktes erfüllen, als Bilder zweier sich senkrecht schneidender Scharen von zu einander parallelen Geradenstücken aufgefasst werden können. Ist jetzt S ein isolierter singulärer innerer Punkt des Netzes, d.h. ist jeder Punkt einer geeignet gewählten Umgebung von S ein regulärer Netzpunkt, so lässt sich S durch eine geschlossene Jordankurve $\mathfrak{p}$ mit zwei Eigenschaften umgeben:[2]

erstens: $\mathfrak{p}$ *begrenzt einen einfach zusammenhängenden Bereich, der ausser* S *nur reguläre Punkte des Netzes enthält*;

zweitens: $\mathfrak{p}$ *setzt sich in seiner Gesamtheit aus endlichen vielen Stücken von Netzbogen zusammen.*

Die Punkte von $\mathfrak{p}$, in denen Stücke verschiedener Netzkurven zusammenstossen, d.h. in denen eine zu $\mathfrak{p}$ gehörige Kurve der ersten Schar sich mit derjenigen Kurve der zweiten Schar schneidet, auf welcher die angrenzende Polygonseite gelegen ist, sollen Ecken des Polygonzuges heissen, und zwar unterscheide man einspringende von ausspringenden Ecken, je nachdem die

[1] H. Hamburger, "Über Kurvennetze mit isolierten Singularitäten auf geschlossenen Flächen." *Math. Zeitschr.* **19** (1924), S. 50-66. Vgl. auch die dort zitierten Arbeiten von Henri Poincaré, Fussnote 8; S. 53 und Ivar Bendixson Fussnote 10, S. 54.

[2] H. Hamburger, *loc. cit.* Fussnote 1; S. 59. Abschnitt **13**.

Fortsetzung der beiden in der Ecke zusammenstossenden Netzbogen über die Ecke hinaus sich beide gleichzeitig ins Innere von $\mathfrak{p}$ hinein fortsetzen lassen oder nicht. Bezeichnet i die Differenz der Anzahl aller ausspringenden Ecken von $\mathfrak{p}$ vermindert um die Anzahl aller einspringenden Ecken von $\mathfrak{p}$, so zeigt sich, dass, obgleich sich ein Polygonzug $\mathfrak{p}$ mit den geforderten zwei Eigenschaften in unendlich mannigfacher Weise konstruieren lässt, alle diese Polygonzüge auf die gleiche Zahl i führen. Mithin ist i eine für die topologische Singularität S invariante charakteristische Zahl, die nunmehr S als Index zugeordnet werde.[3]

3. Erfüllt jetzt das Kurvennetz vollständig eine geschlossene zweiseitige Fläche vom Geschlechte p, wobei alle Flächenpunkte reguläre Netzpunkte sind, ausgenommen k singuläre Punkte S_κ ($\kappa = 1, 2, \cdots, k$), mit zugehörigen Indices i_κ, so gilt die fundamentale Formel

$$\sum_{\kappa=1}^{k}\left(\frac{i_\kappa}{4} - 1\right) = 2p - 2.^{4} \tag{1}$$

4. Man denke sich jetzt auf einer geschlossenen zweiseitigen Fläche vom Geschlecht p mit stetiger Tangentenebene und stetiger Krümmung das System der Krümmungslinien konstruiert; wir erhalten so ein topologisches Kurvennetz, dessen einzige topologischen Singularitäten offenbar die Nabel sind.

Aus Formel (1) folgen jetzt unmittelbar für geschlossene zweiseitige Flächen vom Geschlechte Null die beiden Tatsachen:

I. *Es existiert keine eschlossene Fläche vom Geschlechte* 0 *ganz ohne Nabel.*[5]

Denn bei einer solchen Fläche würde die Summe auf der linken Seite von (1) gleich null werden, während die rechte Seite von (1) gleich -2 ist.

II. *Wenn eine geschlossene Fläche vom Geschlechte* 0 *mit einem einzigen Nabel existieren würde so müsste der Index i dieses Nabels gleich* -4 *sein*; denn für eine solche Fläche folgt aus (1)

$$\frac{i}{4} - 1 = -2, \text{ d.h.} \qquad i = -4.$$

Die Caratheodorysche Vermutung ersetzen wir nunmehr durch die viel weitergehende Behauptung:

$$i \geqq 0^{6}; \tag{2}$$

[3] Über andere Arten, den Index der isolierten Singularität eines Kurvennetzes oder einer einparametrigen Kurvenschar einzuführen, insbesondere über die Definitionen von H. Poincaré, I. Bendixson und Hellmuth Kneser, vgl. den Bericht der in Fussnote 1; angegebenen Arbeit des Verf. S. 53–57. Vgl. auch dort die zahlreichen Fussnoten, die genaure Literaturangaben enthalten.

[4] Ein einfacher Beweis dieser Formel unter sehr allgemeinen Voraussetzungen findet sich l. c. Fussnote 1; S. 58–62.

[5] Ein direkter Beweis (ohne Heranziehung von Formel (1)) des Satzes, dass kein singularitätenfreies Kurvennetz existiert, welches eine Fläche vom Geschlechte 0 vollständig überdeckt, findet sich in der Arbeit des Verf.: "Ein Satz über Kurvennetze auf geschlossenen Flächen." *Sitzber. d. Preuss. Akad. d. Wiss. zu Berlin* **21** (1922), S. 258–262.

[6] Es sei hier eine einfache und anschauliche Definition des Index einer Singularität

denn mit Rücksicht auf die eben formulierte Tatsache II folgt die Caratheodorysche Vermutung unmittelbar aus (2).

5. Um eine möglichst einfache Form für die Differentialgleichung der Krümmungslinien zu erhalten, gehen wir von einer Darstellung der vorgelegten Fläche $\mathfrak{F}$ aus, die wir in Kap. II §5 der vorliegenden Arbeit ableiten werden. Bezeichnen x, y, z die drei rechtwinkligen kartesischen Koordinaten eines Punktes auf $\mathfrak{F}$, so setze man

$$
\begin{aligned}
x &= u - 2W\frac{W_u}{1+W_u^2+W_v^2},\\
y &= v - 2W\frac{W_v}{1+W_u^2+W_v^2},\\
z &= \frac{2W}{1+W_u^2+W_v^2}.
\end{aligned}
\tag{3}
$$

Hierbei sind u, v speziell gewählte Gaussche Flächenparameter, W ist eine dreimal stetig differentiierbare Funktion von u und v; W_u, W_v, W_{uu} usw. bezeichnen in üblicher Weise die partiellen Ableitungen von $W(u, v)$. Durch die Funktion $W(u, v)$ ist somit die Fläche $\mathfrak{F}$ eindeutig bestimmt.

Eine geometrische Deutung dieser Flächendarstellung erhält man, wenn man die Abbildung betrachtet, welche den durch die Formeln (3) bestimmten Flächenpunkt in den Punkt der Ebene mit den rechtwinkligen Koordinaten u, v über führt;[7] dann existiert eine Kugel vom Radius W und dem Mittelpunkt $x = u, y = v, z = W$, welche gleichzeitig $\mathfrak{F}$ im Punkte (3) und die Ebene $z = 0$ im Punkte $x = u, y = v$ berührt, oder anders, noch anschaulicher ausgedrückt: Spiegeln wir das Normalensystem von $\mathfrak{F}$ an den Tangentialebenen der Fläche mit den Koordinaten $x = u, y = v, z = W$, so wird dieses Normalsystem in eine Schar von zur z-Achse parallelen Geraden übergeführt, insbesondere wird

---

für den Fall erwähnt, dass das Kurvennetz, wie z. B. bei den Krümmungslinien durch eine Differentialgleichung 1. Ordnung definiert ist:

Bezeichnet ω den Zuwachs des Winkels einer Integral-kurve, den man bei einem geschlossenen Umlauf um den Nabel erhält, und setzt man $\omega = j\pi$, so besteht zwischen j und i der Zusammenhang $j = 2 - i/2$.

Vgl. z. B. H. Hopf und H. Samelson: "Zum Beweise des Kongruenzsatzes für Eiflächen." *Math. Zeitschr.* **43** (1938), S. 749–766. Siehe insbesondere S. 750.

Der Behauptung (2) $i \geqq 0$ entspricht dann $j \leqq 2$.

Endlich sei noch auf einen Satz von Herrn Schilt über Flachpunkte hingewiesen, in die man ja jeden Nabel durch eine geeignete Inversion (Transformation durch reziproke Radien) überführen kann: Besitzt jeder Flächenpunkt in der Umgebung eines isolierten Flachpunktes negative Krümmung, so ist immer $j \leqq 0$, d.h. somit $i \geqq 2$. Vgl. H. Schilt, "Über die isolierten Nullstellen der Flächenkrümmung und einige Verbiegbarkeitssätze." *Compos. Math.* **5** (1938), S. 239. Siehe insbesondere Nr. 18.

Vgl. auch die oben zitierte Arbeit von H. Hopf und H. Samelson S. 756.

[7] Eine ausführliche Ableitung der Darstellung (3) vermittels Ribancourtransformation findet sich in der Arbeit des Verf. "Ribancourtransformationen und sphärische Abbildung." *Acta Math.* **62** (1933), S. 1–90. Vgl. insbesondere S. 44–46.

hierbei die Flächen-normale im Punkte (3) in eine Gerade verwandelt, die auf der Ebene $z = 0$ im Punkte $x = u, y = v$ senkrecht steht. (vgl. Fig. 4. und die ausführliche Ableitung in Kap. II §5).

6. Wie in Kap. II §6 gezeigt wird, ist die Differentialgleichung der Krümmungslinien der durch die Formeln (3) dargestellten Fläche

$$W_{uv}\,(du^2 - dv^2) + (W_{vv} - W_{uu})\,du\,dv = 0. \tag{4}$$

Diese Differentialgleichung hat vor allem die einfache Eigenschaft, dass, wenn man ihre Integralkurven in der u, v Ebene deutet, diese Kurven ein ebenes orthogonales Netz bilden.

Indem man die Differentialgleichung (4) oder vielmehr eine aus ihr durch Einführung von Polarkoordinaten abgeleitete Differentialgleichung (vgl. Formel (8)) zugrundelegt, lässt sich der Index i einer Singularität von (4) nach einer einfachen allgemeinen Methode bestimmen, welche in Kapitel I auseinandergesetzt ist.

7. Um endlich die Voraussetzungen, unter denen wir den Beweis der Behauptung (2) in Angriff nehmen, genau zu formulieren, substituieren wir in der Funktion $W(u, v)$ die Polarkoordinaten

$$u = \rho \cos\theta, \qquad v = \rho \sin\theta.$$

Wir fordern nunmehr, dass die vorgelegte Fläche $\mathfrak{F}$ durch den Nullpunkt geht, dort einen Nabel hat und die x, y Ebene als Tangentialebene besitzt. Analytisch drücken wir diese Tatsache dadurch aus, dass wir verlangen, W sei in der Umgebung des Nabels von der Form

$$W(\rho, \theta) = \frac{\rho^2}{4R} + \rho^m(w_0(\theta) + F(\rho, \theta)). \tag{5}$$

Hierbei seien $w_0(\theta)$ und $F(\rho, \theta)$ für $0 \leqq \rho \leqq \delta$, $-\infty < \theta < +\infty$ dreimal stetig differentiierbare Funktionen, die in θ offenbar die Periode 2π haben. Ausserdem sei

$$\lim_{\rho \to 0} F(\rho, \theta) \to 0,$$

während $w_0(\theta)$ nicht identisch verschwinden möge; endlich bezeichne m eine positive Zahl > 2, R eine positive Konstante, welche gleich dem Radius der die Fläche $\mathfrak{F}$ im Nabel berührenden Schmiegungskugel ist, während $m - 1$ gleich der Ordnung der Berührung zwischen $\mathfrak{F}$ und der Schmiegungskugel ist.[8]

[8] Der Verlauf der Krümmungslinien in der Umgebung eines Nabelpunktes ist schon mehrfach, u.a. von Cayley, Darboux und E. Picard untersucht worden. Alle diese Autoren behandeln aber nur den Fall $m = 3$ unter einer Reihe weiterer einschränkender Voraussetzungen. Erschöpfend ist der Fall $m = 3$ erst in einer längeren Abhandlung von A. Gullstrand, "Zur Kenntnis der Kreispunkte," *Acta Math.* **29** (1905), S. 59–100 dargestellt worden. Ebenda ist auch der Fall $m = 4$, allerdings unter einschränkenden Voraussetzungen in Angriff genommen; für allgemeines m beschränkt sich Gullstrand nur

Im Teil I Kap. III wird nun, gestützt auf die in Kap. I zur Bestimmung des Index i entwickelte Methode, der Beweis der Behauptung $i \geqq 0$ geführt unter der Voraussetzung, dass

I. *die Funktion $dw_0(\theta)/d\theta$ nur isolierte Nullstellen hat,*

II. *die Funktion $w_0(\theta)$ nur Nullstellen erster und zweiter Ordnung, aber keine Nullstellen dritter und höherer Ordnung besitzt.*

Es gelingt hier, die Abschätzung $i \geqq 0$ sogar in einer noch schärferen Form zu beweisen, nämlich: *Unter den Voraussetzungen I und II lässt sich um den Nabel immer ein aus Stücken von Krümmungslinien zusammengesetzter geschlossener Polygonzug $\mathfrak{p}$ konstruieren, welcher überhaupt keine einspringenden Ecken enthält; ausgenommen ist hierbei der Fall, dass die eine Schar von Krümmungslinien sich spiralförmig um den Nabel als asymptotischem Punkt windet, dann lässt sich der geschlossene Polygonzug $\mathfrak{p}$ so konstruieren, dass er genau eine einspringende und eine ausspringende Ecke enthält.*

Im Teil II wird auf die Voraussetzung II verzichtet, d.h. es werden Funktionen $w_0(\theta)$ zugelassen, welche im Intervall $0 \leqq \theta \leqq 2\pi$ endlich viele Nullstellen beliebig hoher Ordnung besitzen. Unter einigen weiteren Voraussetzungen, die gewiss alle erfüllt sind, wenn nur die Fläche im Nabelpunkt und seiner Umgebung analytisch ist, wird dort die Behauptung $i \geqq 0$ bewiesen werden. Hierzu sind wesentlich tiefer liegende Untersuchungen erforderlich als diejenigen, welche wir hier in Teil I durchzuführen haben.[9])

Kapitel I. Entwicklung einer allgemeinen Methode, um den Index der isolierten Singularität einer Differentialgleichung erster Ordnung zu bestimmen, deren Integralkurven ein ebenes orthogonalfs Kurvensystem bilden

§1. Einführung von Polarkoordinaten. Voraussetzungen über die Koeffizienten der Differentialgleichung

8. In der u, v Ebene sei eine Differentialgleichung

$$X(du^2 - dv^2) + 2Y\,du\,dv = 0 \tag{6}$$

vorgelegt, deren Koeffizienten X, Y in der Umgebung des Nullpunktes der u, v Ebene definiert sind und dort einer Lipschitz-Bedingung genügen. X und Y mögen ausserdem für $u = 0$, $v = 0$ gleichzeitig verschwinden, während sie in einer gewissen Umgebung des Nullpunktes der u, v Ebene keine gemeinsame Nullstelle ausser dem Nullpunkt haben.

---

auf wenige Bemerkungen. Genauere Literaturangaben finden sich bei Darboux, *Leçons sur la théorie générale des surfaces* IV (1896), S. 448, und in der zitierten Arbeit von Gullstrand S. 59 und 60. Über die Bedeutung der in der Darstellung (5) enthaltenen Voraussetzungen über die Fläche $\mathfrak{F}$ verg. Fussnote 16 §7.

[9] Die Überlegungen und Beweise des I. Teiles hat Verf. bereits im Jahre 1923 gefunden und darüber in der Berliner Mathemat. Gesellschaft in der 210. Sitzung am 26. 3. 1924 vorgetragen. Die Untersuchungen zum II. Teil hat Verf. Okt. 1937 begonnen und Sept. 1938 abgeschlossen.

Dann bilden die Integralkurven der Differentialgleichung (6) offenbar in der Umgebung des Nullpunktes der u, v Ebene ein orthogonales Kurvensystem mit dem Nullpunkte als einzige Singularität.

9. Wir führen Polarkoordinaten ein

$$u = \rho \cos \theta, \qquad v = \rho \sin \theta. \tag{7}$$

Dann wird

$$du^2 - dv^2 = \cos 2\theta \cdot (d\rho^2 - \rho^2 d\theta^2) - 2 \sin 2\theta \cdot \rho \, d\rho \, d\theta.$$

$$2 \, du \, dv = \sin 2\theta \cdot (d\rho^2 - \rho^2 d\theta^2) + 2 \cos 2\theta \cdot \rho \, d\rho \, d\theta,$$

und für (6) erhält man

$$A(d\rho^2 - \rho^2 d\theta^2) + 2B\rho \, d\rho \, d\theta = 0, \tag{8}$$

wobei

$$\begin{cases} A = X \cos 2\theta + Y \sin 2\theta, \\ B = -X \sin 2\theta + Y \cos 2\theta \end{cases} \tag{9}$$

gesetzt ist.

10. Aus den Voraussetzungen für X und Y folgt für A und B:

1. *A und B verschwinden beide für $\rho = 0$;*
2. *in einer gewissen einseitigen Umgebung $\mathfrak{U}$ der Geraden $\rho = 0$ in der $\rho \cdot \theta$-Ebene ($\rho > 0$) konnen A und B nicht gleichzeitig verschwinden;*
3. *A und B haben beide in θ die Periode 2π;*
4. *A und B genügen in $\mathfrak{U}$ einer Lipschitz-Bedingung.*

Aus der Voraussetzung 4. folgt unmittelbar, dass für $\rho > 0$ durch jeden Punkt von $\mathfrak{U}$ zwei und nur zwei Integralkurven laufen, welche der Differentialgleichung (8) genügen. Weiter setzen wir für A und B voraus:

5. *Die Umgebung $\mathfrak{U}$ der Geraden $\rho = 0$ lasse sich so klein wählen, dass der Koeffizient A dort nur auf isolierten stetigen Kurven $\mathfrak{c}_\nu$ verschwindet, welche sämtlich einen bestimmten Punkt $\theta = \theta_\nu$ der Geraden $\rho = 0$ mit einem Punkte des Randes von $\mathfrak{U}$ verbinden, derart dass zwei beliebige Kurven $\mathfrak{c}_\nu$ und $\mathfrak{c}_\mu$ höchstens auf der Geraden $\rho = 0$, aber nicht im Innern von $\mathfrak{U}$, einen gemeinsamen Schnittpunkt haben. Endlich sei N die Anzahl der Kurven $\mathfrak{c}_\nu$ ($\nu = 1, 2, \cdots, N$), deren Anfangspunkte $\theta = \theta_\nu$ auf der Geraden $\rho = 0$ sämtlich im Periodenstreifen $0 \leqq \theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_\nu \leqq \theta_{\nu+1} \leqq \cdots \leqq \theta_N < 2\pi$ gelegen sind.*

§2. Geometrische Betrachtungen über die Integralkurven in der Nähe der Kurven $\mathfrak{c}_\nu$

11. Die Auflösung der quadratischen Gleichung (8) ergibt

$$\frac{d\rho}{d\theta} = -\rho \frac{B}{A} \left(1 \mp \left(1 + \frac{A^2}{B^2}\right)^{\frac{1}{2}}\right).$$

In einer hinreichenden Nähe der in Voraussetzung 5. definierten Kurven c_ν lässt sich, solange $A/B < 1$, die Quadratwurzel in eine konvergente Potenzreihe entwickeln, und man erhält

$$\left(\frac{d\rho}{d\theta}\right)_1 = \rho \sum_{\nu=1}^{\infty} \binom{\frac{1}{2}}{\nu} \frac{A^{2\nu-1}}{B^{2\nu-1}}, \tag{10}$$

$$\left(\frac{d\rho}{d\theta}\right)_2 = -2\rho \frac{B}{A} - \rho \sum_{\nu=1}^{\infty} \binom{\frac{1}{2}}{\nu} \frac{A^{2\nu-1}}{B^{2\nu-1}}. \tag{11}$$

Hieraus folgt, dass für das Vorzeichen, wenn A hinreichend klein bleibt, das erste Glied der Entwicklung den Ausschlag gibt:

$$\operatorname{sg}\left(\frac{d\rho}{d\theta}\right)_1 = \operatorname{sg}\frac{A}{B}, \qquad \operatorname{sg}\left(\frac{d\rho}{d\theta}\right)_2 = -\operatorname{sg}\frac{B}{A} = -\operatorname{sg}\left(\frac{d\rho}{d\theta}\right)_1. \tag{12}$$

12. Nunmehr unterscheiden wir drei Arten von Kurven: Wir sagen, c_ν *ist von erster Art* (kurz eine Kurve c_ν'), *wenn A/B beim Überschreiten von c_ν' mit wachsendem θ von negativen zu positiven Werten übergeht,*

c_ν *is von zweiter Art* (kurz eine Kurve c_ν''), *wenn A/B beim Überschreiten von c_ν'' mit wachsendem θ von positiven zu negativen Werten übergeht,*

c_ν *ist von dritter Art* (kurz eine Kurve c_ν'''), *wenn A/B auf beiden Seiten von c_ν''' das gleiche Vorzeichen hat.*

Wegen (12) hat offenbar diejenige Integralkurve von (8), deren Tangente gleich $(d\rho/d\theta)_1$ ist, in ihrem Schnittpunkt mit c_ν

ein Minimum, wenn c_ν erster Art ist,

ein Maximum, wenn c_ν zweiter Art ist,

einen Wendepunkt, wenn c_ν dritter Art ist.

13. Endlich zeigt die Differentialgleichung (8) unmittelbar, dass eine Integralkurve, für $\rho > 0$, ganz gleich ob ihre Tangente $(d\rho/d\theta)_1$ oder $(d\rho/d\theta)_2$ ist, das Vorzeichen ihrer Tangente nur in ihren Schnittpunkten mit den Kurven c_ν' oder c_ν'' wechseln kann, da $(d\rho/d\theta)_1$ und $(d\rho/d\theta)_2$ nur in solchen Punkten 0 bzw. ∞ werden, wo $A = 0$ ist.

§3. Beschreibung der Methode zur Bestimmung des Index

14. Von den in Voraussetzung 5. §1 beschriebenen N Kurven c_ν lassen wir jetzt alle Kurven 3. Art. fort. Dann mögen noch n Kurven c_ν ($\nu = 1, 2 \cdots n$) übrig bleiben, die sämtlich 1. Art oder 2. Art sind, und die wir uns so geordnet denken, dass die Kurve $c_{\nu+1}$ ganz rechts von c_ν gelegen ist. Ausserdem erhält man wegen der Periodizitätseigenschaft von A (Voraussetzung 3. §1) die Kurve c_{n+1}, indem man c_1 parallel zu sich in der Richtung der θ-Achse um 2π verschiebt.

Wir betrachten nun den Bereich $\mathfrak{B}_\nu$, der von c_ν, $c_{\nu+1}$ dem Stück des Randes von $\mathfrak{U}$ zwischen c_ν und $c_{\nu+1}$ und dem Stück $\theta_\nu \leqq \theta \leqq \theta_{\nu+1}$ (vgl. Voraussetzung 5. §1) der Geraden $\rho = 0$ begrenzt wird. Falls $\theta_\nu = \theta_{\nu+1}$ zieht sich dieses Geradenstück auf einen Punkt zusammen. Dann kann die Tangente einer beliebigen Integralkurve nach der Bemerkung aus Abschnitt **13.** ihr Vorzeichen im Innern

von $\mathfrak{B}_\nu$ nirgends wechseln, da zwischen c_ν und $c_{\nu+1}$ keine weitere Kurve 1. Art oder 2. Art gelegen ist.

Wir wollen jetzt in $\mathfrak{B}_\nu$ den in der Einleitung Abschnitt **2** beschriebenen Polygonzug $\mathfrak{p}$ konstruieren, der sich aus Integralkurvenstücken zusammensetzt und dann zur Bestimmung des Index i führt, wobei wir drei Fälle zu unterscheiden haben.

15. 1. FALL. c_ν *und* $c_{\nu+1}$ *sind beide von* 1. *Art; wir nennen dann auch den zugehörigen Bereich* $\mathfrak{B}_\nu$ *von* 1. *Art.* Wir betrachten das Netz der Integralkurven von (8) in $\mathfrak{B}_\nu$. Dieses Kurvennetz kann man in zwei Scharen aufteilen, und zwar rechnen wir eine Integralkurve zur Schar I oder II, je nachdem ihre Tangente $d\rho/d\theta \geqq 0$ oder $\leqq 0$ ist. Die Grenzfälle $d\rho/d\theta = 0$ bzw. ∞ können auch eintreten, da ja $\mathfrak{B}_\nu$ noch Kurven c''' im Innern enthalten kann. Offenbar können sich zwei Integralkurven der gleichen Schar wegen Voraussetzung 4. aus §1 in $\mathfrak{B}_\nu$ nirgends schneiden.

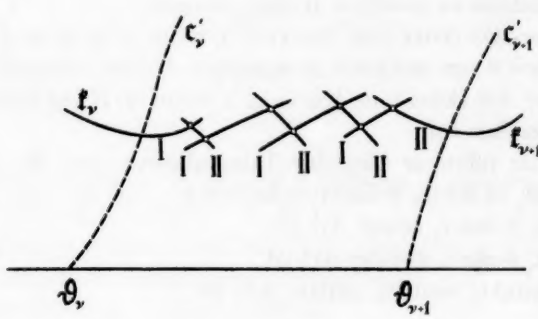

FIG. 1

Wir betrachten nunmehr ein Integralkurvenstück $\mathfrak{k}_\nu$, das c'_ν schneidet und dessen Tangente in der Umgebung von c_ν gleich $(d\rho/d\theta)_1$ wird (vgl. Formel (10)). Dann gehört dieses Kurvenstück zur Schar I, da wegen (12) rechts von c'_ν die Tangente $(d\rho/d\theta)_1 > 0$ wird. Ebenso erkennen wir, dass ein Integralkurvenstück $\mathfrak{k}_{\nu+1}$ das $c'_{\nu+1}$ mit der Tangente $(d\rho/d\theta)_1$ schneidet, zur Schar II gehört, da wegen (12) links von $c'_{\nu+1}$ die Tangente $(d\rho/d\theta)_1 < 0$ wird.

Eine leichte Überlegung zeigt uns,[10] dass sich jetzt ein Polygonzug konstruieren lässt, der sich aus einer endlichen geraden Anzahl von Integralkurvenstücken zusammensetzt und einen Punkt von $\mathfrak{k}_\nu$ mit einem Punkt von $\mathfrak{k}_{\nu+1}$ verbindet (vgl. Fig. 1.). In diesem Polygonzug werden offenbar die Kurvenstücke der Scharen I und II dauernd abwechseln. Ferner wird man beim Übergang von der Schar I zur Schar II eine ausspringende, beim Übergang von der Schar II zur Schar I eine einspringende Ecke erhalten, sodass in dem konstruierten Polygonzug die ausspringenden und einspringenden Ecken immer alternieren.

[10] Vgl. l. c. Fussnote 1; S. 59 Abschn. **13.**

Da $\mathfrak{k}_\nu$ zu I und $\mathfrak{k}_{\nu+1}$ zu II gehört, wird die Folge der Ecken mit einer ausspringenden Ecke beginnen und einer ausspringenden Ecke aufhören. Die Anzahl der ausspringenden Ecken des Polygonzuges, der sich von einem Punkte von $\mathfrak{c}'_\nu$ bis zu einem Punkte von $\mathfrak{c}'_{\nu+1}$ erstreckt, ist somit genau um 1 grösser als die Anzahl der einspringenden Ecken, d.h. aber, dass der Polygonzug in einem Bereiche $\mathfrak{B}_\nu$ von 1. Art für den Index i den Beitrag $+1$ liefert.

16. 2. FALL. $\mathfrak{c}_\nu$ *und* $\mathfrak{c}_{\nu+1}$ *sind beide von* 2. *Art; wir nennen dann auch den zugehörigen Bereich* $\mathfrak{B}_\nu$ *von* 2. *Art.* Wir unterscheiden ebenso wie beim 1. Fall die Integralkurven der Schar I und der Schar II und bemerken, dass jetzt das Integralkurvenstück $\mathfrak{k}_\nu$, das $\mathfrak{c}''_\nu$ schneidet und dessen Tangente in der Umgebung von $\mathfrak{c}''_\nu$ gleich $(d\rho/d\theta)_1$ ist, zur Schar II gehört, da rechts von $\mathfrak{c}''_\nu$ wegen (12) $(d\rho/d\theta)_1 < 0$ wird, während $\mathfrak{k}_{\nu+1}$ das $\mathfrak{c}''_{\nu+1}$ mit der Tangente $(d\rho/d\theta)_1$ schneidet, zur Schar I gehört. da links von $\mathfrak{c}''_{\nu+1}$ wegen (12) $(d\rho/d\theta)_1 > 0$ wird.

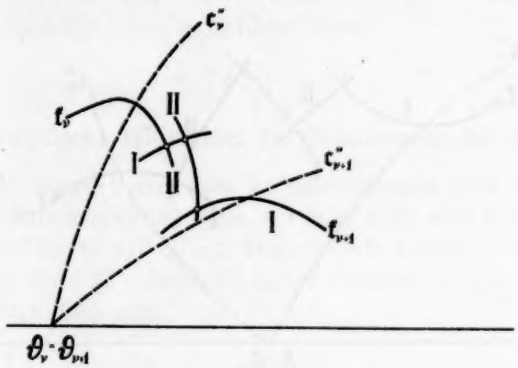

FIG. 2

Konstruieren wir jetzt ebenso wie im 1. Fall einen Polygonzug, der sich aus endlich vielen Integralkurvenstücken zusammensetzt und einen Punkt von $\mathfrak{k}_\nu$ mit einem Punkt von $\mathfrak{k}_{\nu+1}$ verbindet, so werden auch hier im Polygonzuge einerseits die Kurvenstücke der Schar I und der Schar II, anderseits die einspringenden und ausspringenden Ecken dauernd wechseln (vgl. Fig. 2). Da $\mathfrak{k}_\nu$ zur Schar II, $\mathfrak{k}_{\nu+1}$ zur Schar I gehört, so wird die Folge der Ecken des Polygonzuges mit einer einspringenden Ecke beginnen und einer einspringenden Ecke aufhören. Die Anzahl der einspringenden Ecken des Polygonzuges, der sich von einem Punkte von $\mathfrak{c}''_\nu$ zu einem Punkte von $\mathfrak{c}''_{\nu+1}$ erstreckt, ist somit genau um 1 grösser als die Anzahl der ausspringenden Ecken, d.h. aber, dass der Polygonzug in einem Bereiche $\mathfrak{B}_\nu$ von 2. Art für den Index i den Beitrag -1 liefert.

17. 3. FALL. *Von den beiden Kurven* $\mathfrak{c}_\nu$ *und* $\mathfrak{c}_{\nu+1}$ *ist die eine von* 1. *Art, die andere von* 2. *Art. Dann sagen wir, der zugehörige Bereich* $\mathfrak{B}_\nu$ *sei von* 3. *Art.* In diesem Fall müssen die beiden Kurvenstücke $\mathfrak{k}_\nu$ und $\mathfrak{k}_{\nu+1}$, die $\mathfrak{c}_\nu$ bzw. $\mathfrak{c}_{\nu+1}$ mit

der Tangente $(d\rho/d\theta)_1$ schneiden, beide zur gleichen Schar ghören und zwar offenbar beide zur Schar I, wenn $\mathfrak{c}_\nu$ von 1. Art, $\mathfrak{c}_{\nu+1}$ von 2. Art,—beide zur Schar II, wenn $\mathfrak{c}_\nu$ von 2. Art, $\mathfrak{c}_{\nu+1}$ von 1. Art ist.

Konstruiert man jetzt den Polygonzug aus Integralkurvenstücken, der einen Punkt von $\mathfrak{k}_\nu$ mit einem Punkt von $\mathfrak{k}_{\nu+1}$ verbindet, so wird im 3. Falle, da ja auch hier wieder im Polygonzug einerseits die Kurvenstücke der Schar I und der Schar II, anderseits aber auch die einspringenden und ausspringenden Ecken dauernd wechseln, offenbar die Anzahl der Kurvenstücke des Polygonzuges ungerade und damit die Anzahl der Ecken gerade sein (vgl. Fig 3). D.h. aber der Polygonzug enthält ebenso viel ausspringende wie einspringende Ecken. Somit liefert der Polygonzug in einem Bereich $\mathfrak{B}_\nu$ von 3. Art zum Index i den Beitrag 0.

18. Nunmehr konstruiere man für jedes $\nu(\nu = 1, 2, \cdots, n, n+1)$ ein Integralkurvenstück $\mathfrak{k}_\nu$, das $\mathfrak{c}_\nu$ mit der Tangente $(d\rho/d\theta)_1$ schneidet, wobei man wegen

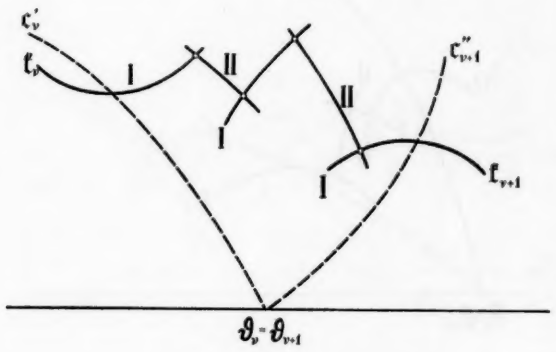

FIG. 3

Voraussetzung 3 §1. $\mathfrak{k}_{n+1}$ aus $\mathfrak{k}_1$ erhält, indem man $\mathfrak{k}_1$ parallel zu sich in der Richtung der θ Achse um 2π verschiebt, und konstruiere dann weiter in jedem der n Bereiche $\mathfrak{B}_\nu$ $(\nu = 1, 2, \cdots, n)$ einen Polygonzug in der oben angegebenen Weise, je nachdem $\mathfrak{B}_\nu$ ein Bereich 1., 2. oder 3. Art ist. Auf diese Weise gelangt man zu einem zusammenhängenden Polygonzug $\mathfrak{p}$ der ρ, θ Ebene, der einen Punkt von $\mathfrak{k}_1$ mit einem Punkt von $\mathfrak{k}_{n+1}$ verbindet. Geht man jetzt von der ρ, θ Ebene wieder zu der u, v Ebene über, indem man die zu (7) inverse Transformation ausführt, so wird aus $\mathfrak{p}$ ein geschlossener Polygonzug $\mathfrak{p}^*$ in der u, v Ebene, da nach der Transformation das Bild von $\mathfrak{k}_{n+1}$ mit dem Bild von $\mathfrak{k}_1$ zusammenfällt. Ferner gehen die Ecken von $\mathfrak{p}$ in die Ecken von $\mathfrak{p}^*$ über und behalten offenbar nach der Transformation ihre Eigenschaft bei, ausspringend oder einspringend zu sein.

19. Der Gesamtindex i des Nabels setzt sich mithin zusammen aus den einzelnen Beiträgen, den jeder der n Bereiche $\mathfrak{B}_\nu$ zu dem Index i liefert.

Berücksichtigt man die Ergenbnisse der drei Abschnitte **15.**, **16.**, **17.**, so ergibt sich der Satz:

Der Index i des Nabels ist gleich der Differenz aus der Anzahl der Bereiche $\mathfrak{B}_\nu$ von 1. Art vermindert um die Anzahl der Bereiche $\mathfrak{B}_\nu$ von 2. Art.

Es bezeichne jetzt

$$(13)\qquad \begin{cases} \{1\} \text{ die Anzahl der Bereiche } \mathfrak{B}_\nu \text{ von 1. Art,} \\ \{2\} \text{ die Anzahl der Bereiche } \mathfrak{B}_\nu \text{ von 2. Art,} \\ \{+\} \text{ die Anzahl der Kurven } \mathfrak{c}_\nu \text{ von 1. Art,} \\ \{-\} \text{ die Anzahl der Kurven } \mathfrak{c}_\nu \text{ von 2. Art,} \end{cases}$$

dann ist

$$i = \{1\} - \{2\}.$$

Nun zeigt aber eine leichte Überlegung, dass

$$\{1\} - \{2\} = \{+\} - \{-\}$$

ist. Somit erhält man die wichtige Schlussformel

$$(14)\qquad i = \{+\} - \{-\}.$$

§4. Analytische Hilfsmittel zur Bestimmung des Index

20. Zu den in §1 angegebenen fünf Voraussetzungen über die Koeffizienten A, B der Differentialgleichung (8) fügen wir jetzt noch zwei weitere hinzu.

6. *Für hinreichend kleine* ρ $(0 \leqq \rho \leqq \delta)$ *werde jede Kurve* $\mathfrak{c}_\nu$, *auch die Kurven 3. Art* (vgl. Voraussetzung 5 §1), *durch die stetige Funktion* $\theta = \Theta_\nu(\rho)(\nu = 1, 2, \cdots, N, N+1)$ *analytisch dargestellt.*

Dann ist

$$\Theta_{\nu+1}(\rho) > \Theta_\nu(\rho) \text{ für } \rho > 0,$$

$$\Theta_{N+1}(\rho) = \Theta_1(\rho) + 2\pi.$$

7. *Für jedes* $\nu(\nu = 1, 2, \cdots, N)$ *lasse sich* A *in das Produkt*

$$(15)\qquad A(\rho, \theta) = (\theta - \Theta_\nu(\rho))^{k_\nu} A_\nu(\rho, \theta)$$

zerlegen, wobei k_ν *eine ganze positive Zahl bedeutet, und* $A_\nu(\rho, \Theta_\nu(\rho)) \neq 0$ *für* $0 < \rho \leqq \delta$ *ist.*

Dann ist offenbar:

$\mathfrak{c}_\nu$ von 1. Art, wenn k_ν ungerade und

$$(16)\qquad \operatorname{sg} \frac{A_\nu(\rho, \Theta_\nu(\rho))}{B(\rho, \Theta_\nu(\rho))} = +1,$$

$\mathfrak{c}_\nu$ von 2. Art, wenn k_ν ungerade und

$$(16')\qquad \operatorname{sg} \frac{A_\nu(\rho, \Theta_\nu(\rho))}{B(\rho, \Theta_\nu(\rho))} = -1,$$

$\mathfrak{c}_\nu$ von 3. Art, wenn k_ν gerade ist.

Bezeichnet k die grösste der N Zahlen k_ν und ist $A(\rho, \theta)$ k mal partiell nach θ stetig differentiierbar, so wird

$$A_\nu(\rho, \Theta_\nu(\rho)) = \frac{1}{k_\nu!} \frac{\partial^{k_\nu} A(\rho, \theta)}{\partial \theta^{k_\nu}} \Bigg| \theta = \Theta_\nu(\rho). \tag{17}$$

Mit Rücksicht auf die Formeln (16), (16′) und (17) lässt sich der Index i statt durch den symbolischen Ausdruck (14) auch durch die Summe

$$i = \sum_\nu{}' \, sg \, \frac{\dfrac{\partial^{k_\nu} A(\rho, \theta)}{\partial \theta^{k_\nu}}}{2B(\rho, \theta)} \Bigg| \theta = \Theta_\nu(\rho) \tag{18}$$

darstellen; hierbei ist die Summation nur über diejenigen der N Kurven $\theta = \Theta_\nu(\rho)$ zu erstrecken, welche von ungeradem Vielfachheitsgrad k_ν sind.

Die Formel (18) wird den Ausgangspunkt unserer Untersuchungen im Teil II bilden, wo wir weitergehende analytishce Voraussetzungen als hier im Teil I machen werden.

Kapitel II. Ableitung der Differentialgleichung der Krümmungslinien

§5. Eine spezielle Darstellung der Fläche

21. Die vorgelegte Fläche mit stetiger Tangentenebene und stetiger Krümmung denken wir uns durch einen Vektor $\mathfrak{x}$ bestimmt, dessen drei zu einander rechtwinklige Komponenten x, y, z als dreimal stetig differentiierbare Funktionen zweier Veränderlichen α, β gegeben seien. Hält man den Anfangspunkt des Vektors $\mathfrak{x}$ fest, so beschreibe der Endpunkt die vorgelegte Fläche, die wir kurz die Fläche $\mathfrak{x}$ nennen wollen. Diese Fläche enthalte den Nullpunkt und habe dort die x, y-Ebene zur Tangentenebene. $\mathfrak{e}$ sei der zur positiven z-Achse parallele Einheitsvektor, $\mathfrak{u}$ sei der Normaleinheinsvektor der Fläche $\mathfrak{x}$; sein Vorzeichen sei so festgelegt, dass im Nullpunkt $\mathfrak{u} = -\mathfrak{e}$ wird.

Nunmehr bilde man die Fläche $\mathfrak{x}$ vermittels der Formel

$$\mathfrak{X} = \mathfrak{x} - \frac{\mathfrak{x}\mathfrak{e}}{1 - \mathfrak{u}\mathfrak{e}} (\mathfrak{e} - \mathfrak{u}). \tag{19}$$

auf die x, y-Ebene ab.[11]

Aus Formel (19) folgt nämlich unmittelbar $\mathfrak{e}\mathfrak{X} = 0$ d.h. aber dass der Vektor $\mathfrak{X}$ der x, y-Ebene parallel ist oder aber, wenn der Anfangspunkt des Vektors $\mathfrak{X}$ im Nullpunkt festgehalten wird, dass sein Endpunkt ein Gebiet der x, y-Ebene beschreibt. u, v seien die beiden rechtwinkligen Koordinaten dieses Endpunktes oder, was dasselbe ist, die beiden rechtwinkligen Komponenten von $\mathfrak{X}$, sodass also durch (19) der Flächenpunkt mit den Koordinaten x, y, z in einem Punkt der x, y-Ebene mit den Koordinaten u, v übergeführt wird.

[11] Vgl. l. c. Fussnote 7; S. 44 Formel (106).

Um eine einfache geometrische Deutung der Transformation (19) zu erhalten, setze man die Grösse

$$\frac{\mathfrak{x}\mathfrak{e}}{1 - \mathfrak{u}\mathfrak{e}} = W. \tag{20}$$

Dann ergibt sich aus (19)

$$\mathfrak{X} + W\mathfrak{e} = \mathfrak{x} + W\mathfrak{u}. \tag{21}$$

Bezeichnet man den in (21) definierten Vektor mit $\mathfrak{s}$, so folgt nach leichter Rechnung aus (21)

$$ed\mathfrak{s} = \mathfrak{u}d\mathfrak{s} = dW, \tag{22}$$

und jetzt erkennt man unmittelbar aus (21) und (22), dass die Normale $\mathfrak{x} + t\mathfrak{u}$ der Fläche $\mathfrak{x}$ durch Spiegelung an der Tangentialebene der Fläche $\mathfrak{s}$ in die auf der x,y-Ebene senkrechten Gerade $\mathfrak{X} + t\mathfrak{e}$ übergeführt wird (vgl. Fig. 4.). W

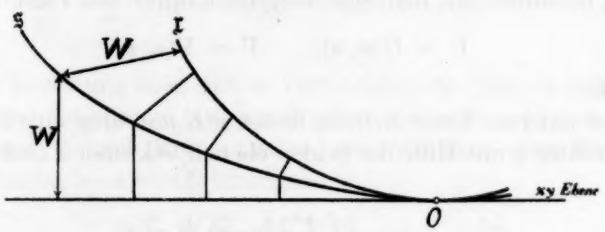

FIG. 4

wird gleich der Entfernung desjenigen Punktes auf $\mathfrak{s}$, in welchem der zu spiegelnde Strahl einfällt, einerseits von der x,y-Ebene, andrerseits von der Fläche $\mathfrak{x}$.

22. Die bisherigen Überlegungen lassen erkennen, wie sich, wenn die Fläche $\mathfrak{x}$ vorgelegt ist, W als Funktion von u and v aus den Beziehungen (19) und (20) ergibt. Jetzt sollen die Formeln (3) der Einleitung abgeleitet werden, welche zeigen, wie man die Fläche $\mathfrak{x}$ bestimmt, wenn W als dreimal stetig differentiierbare Funktion von u und v gegeben ist.

Zu diesem Zwecke bilde man die von dem Normalvektor $\mathfrak{u}$ beschriebene Einheitskugel durch stereographische Projektion auf die x,y-Ebene ab. Dies geschieht vermittels der Relation[12]

$$\mathfrak{U} = \mathfrak{u} - \frac{\mathfrak{u}\mathfrak{e}}{1 - \mathfrak{u}\mathfrak{e}}(\mathfrak{e} - \mathfrak{u}). \tag{23}$$

Hierbei ist $\mathfrak{U}$ ein offenbar zur x,y-Ebene paralleler Vektor, dessen Endpunkt die stereographische Projektion des Vektors $\mathfrak{u}$ beschreibt, wenn der Anfangspunkt

[12] Vgl. l. c. Fussnote 7; S. 44 Formel (107).

des Vektors im Nullpunkt festgehalten wird. Umgekehrt wird $\mathfrak{u}$ aus $\mathfrak{U}$ durch die Formel[13]

$$\mathfrak{u} - \mathfrak{e} = 2\frac{\mathfrak{U} - \mathfrak{e}}{1 + \mathfrak{U}^2} \tag{24}$$

bestimmt.

Jedem Punkt $\mathfrak{x}$ der vorgelegten Fläche sind somit jetzt zwei Punkte der x,y-Ebene zugeordnet: erstens der Endpunkt des Vektors $\mathfrak{X}$ mit den Koordinaten u, v durch die Relation (19), zweitens der Endpunkt des Vektors $\mathfrak{U}$, dessen rechtwinklige Koordinaten wir mit U, V bezeichnen wollen, durch Vermittlung der stereographischen Projektion des der Fläche $\mathfrak{x}$ zugeordneten Normaleinheitsvektors $\mathfrak{u}$ nach Formel (23). Diese zwei Abbildungen $\mathfrak{x} \to \mathfrak{X}$ und $\mathfrak{x} \to \mathfrak{U}$ sind aber dann sicher beide umkehrbar eindeutig, wenn man sich auf eine hinreichend kleine Umgebung eines konvexen Flächenpunktes $\mathfrak{x}$ beschränkt. Hierdurch ist dann aber weiter eine umkehrbar eindeutige Abbildung der x,y-Ebene in sich, $\mathfrak{X} \to \mathfrak{U}$, bestimmt, die man sich analytisch durch das Funktionenpaar

$$U = U(u, v), \qquad V = V(u, v)$$

gegeben denke.

In der oben zitierten Untersuchung findet sich nun aber eine Formel, welche den Flächenvektor $\mathfrak{x}$ mit Hilfe der beiden ebenen Vektoren $\mathfrak{X}$ und $\mathfrak{U}$ ausdrückt[14], und zwar ist

$$\mathfrak{x} = \mathfrak{X} - 2\left(\int_0^{u,v} \mathfrak{U}\, d\mathfrak{X}\right)\frac{\mathfrak{U} - \mathfrak{e}}{1 + \mathfrak{U}^2}, \tag{25}$$

wobei sich zeigen lässt, dass

$$\mathfrak{U} d\mathfrak{X} = U du + V dv$$

ein vollständiges Differential ist.

Zieht man noch Formel (24) heran, so erhält man

$$\mathfrak{x} = \mathfrak{X} - \left(\int_0^{u,v} \mathfrak{U}\, d\mathfrak{X}\right)(\mathfrak{u} - \mathfrak{e}).$$

und nunmehr lässt ein Vergleich mit Formel (19) und (20) erkennen. dass

$$\int_0^{u,v} \mathfrak{U}\, d\mathfrak{X} = W$$

ist, d.h. aber, es ist

$$U = W_u\,; \qquad V = W_v. \tag{26}$$

Setzt man die so erhaltenen Werte für $\int_0^{u,v} \mathfrak{U} d\mathfrak{X}$ und für die Komponenten U, V

[13] Vgl. l. c. Fussnote 7; S. 45 Formel (108) und (109).
[14] Vgl. l. c. Fussnote 7; S. 47 Formel (112).

von $\mathfrak{U}$ in Formel (25) ein und schreibt Formel (25) für die drei Komponenten x, y, z von $\mathfrak{x}$ einzeln auf, so ergeben sich die Formeln (3) der Einhleitung, welche zeigen, wie sich die Fläche $\mathfrak{x}$ vermittels der Funktion $W(u, v)$ darstellen lässt.

§6. Ableitung der Differentialgleichung der Krümmungslinien

23. In einem gewöhnlichen Punkt der Fläche $\mathfrak{x}$, in dem die beiden Hauptkrümmungsrichtungen bestimmt sind, seien für den Augenblick die beiden infinitesimalen Vektoren in Richtung der Krümmungslinien mit $d_1\mathfrak{x}$, $d_2\mathfrak{x}$ bezeichnet. Diese beiden Richtungen werden durch die Abbildung (19) in die beiden infinitesimalen ebenen Vektoren $d_1\mathfrak{X}$, $d_2\mathfrak{X}$, durch die Abbildung (23) in in $d_1\mathfrak{U}$, $d_2\mathfrak{U}$ übergeführt.

Nun gibt es aber einen Satz[15], nach dem

ERSTENS: $d_1\mathfrak{X} \cdot d_2\mathfrak{X} = 0$, $d_1\mathfrak{U} \cdot d_2\mathfrak{U} = 0$;

ZWEITENS: *die Richtung* $d_1\mathfrak{X}$ *zu* $d_1\mathfrak{U}$, *und die Richtung* $d_2\mathfrak{X}$ *zu* $d_2\mathfrak{U}$ *parallel ist*; *d.h. aber es existieren zwei Funktionen* $\lambda_1(u, v)$, $\lambda_2(u, v)$, *derart dass*

$$d_1\mathfrak{U} = \lambda_1 d_1\mathfrak{X}, \qquad d_2\mathfrak{U} = \lambda_2 d_2\mathfrak{X} \tag{27}$$

ist.

Aus dieser Beziehung lässt sich in Verbindung mit (26) die Differentialgleichung der Krümmungslinien fast unmittelbar herleiten. Trennt man nämlich die Gleichung (27) nach Komponenten, so erhält man etwa für die Richtung d_1u: d_1v die beiden linearen Gleichungen

$$d_1W_u = W_{uu}d_1u + W_{uv}d_1v = \lambda_1 d_1u,$$
$$d_1W_v = W_{vu}d_1u + W_{vv}d_1v = \lambda_1 d_1v,$$

und die Elimination von λ_1 ergibt—indem wir jetzt einfach wieder du, dv statt d_1u, d_1v schreiben—die in Formel (4) der Einleitung angegebene Form für die Differentialgleichung der Krümmungslinien:

$$W_{uv}(du^2 - dv^2) + (W_{vv} - W_{uu})dudv = 0.$$

24. Nunmehr hat man noch, um die in Kap. I beschriebene Methode anwenden zu können, wie in Abschnitt **9.** §1 Polarkoordinaten einzuführen. Es werde wieder gesetzt

$$u = \rho \cos\theta, \qquad v = \rho \sin\theta. \tag{28}$$

Dann wird, wie eine leichte Rechnung zeigt,

$$\begin{cases} W_u = W_\rho \cos\theta - \dfrac{W_\theta}{\rho}\sin\theta, \\ W_v = W_\rho \sin\theta + \dfrac{W_\theta}{\rho}\cos\theta, \end{cases} \tag{29}$$

[15] Vgl. l. c. Fussnote 7; S. 44 die mit I und II bezeichneten Sätze.

$$(30)\quad \begin{cases} W_{uv} = -\left(\dfrac{W_\rho}{\rho}\dfrac{\sin 2\theta}{2} + \dfrac{W_\theta}{\rho^2}\cos 2\theta\right) \\ \qquad\qquad + \left(W_{\rho\rho} - \dfrac{W_{\theta\theta}}{\rho^2}\right)\dfrac{\sin 2\theta}{2} + \dfrac{W_{\rho\theta}}{\rho}\cos 2\theta, \\ \dfrac{W_{vv} - W_{uu}}{2} = \dfrac{W_\rho}{\rho}\dfrac{\cos 2\theta}{2} - \dfrac{W_\theta}{\rho^2}\sin 2\theta \\ \qquad\qquad - \left(W_{\rho\rho} - \dfrac{W_{\theta\theta}}{\rho^2}\right)\dfrac{\cos 2\theta}{2} + \dfrac{W_{\rho\theta}}{\rho}\sin 2\theta. \end{cases}$$

Die Transformation (28) führt, wie in Abschnitt **9.** §1 gezeigt wurde, die Differentialgleichung (4) in die Form (8)

$$A(d\rho^2 - \rho^2 d\theta^2) + 2B\rho d\rho d\theta = 0$$

über, wobei man die Ausdrücke für die Koeffizienten A und B erhält, wenn man in die Formeln (9) für X bzw. Y die Ausdrücke (30) für W_{uv} bzw. $\frac{1}{2}(W_{vv} - W_{uu})$ einsetzt. Diese Rechnung ergibt schliesslich

$$(31)\quad \begin{cases} A = \dfrac{W_{\rho\theta}}{\rho} - \dfrac{W_\theta}{\rho^2}, \\ B = \dfrac{1}{2}\left(\dfrac{W_{\theta\theta}}{\rho^2} - W_{\rho\rho} + \dfrac{W_\rho}{\rho}\right). \end{cases}$$

Will man endlich auch in die Formeln (3) der Einleitung Polarkoordinaten einführen, so erhält man wegen (29) die für manche Zwecke wichtige Darstellung der Koordinaten des Flächenpunktes durch die Funktion $W(\rho, \theta)$; und zwar wird

$$(32)\quad \begin{cases} x = \rho\cos\theta - 2W\dfrac{W_\rho\cos\theta - \dfrac{W_\theta}{\rho}\sin\theta}{1 + W_\rho^2 + \dfrac{W_\theta^2}{\rho^2}}, \\ y = \rho\sin\theta - 2W\dfrac{W_\rho\sin\theta + \dfrac{W_\theta}{\rho}\cos\theta}{1 + W_\rho^2 + \dfrac{W_\theta^2}{\rho^2}}, \\ z = \dfrac{2W}{1 + W_\rho^2 + \dfrac{W_\theta^2}{\rho^2}}. \end{cases}$$

Kapitel III. Die Bestimmung des Nabelindex

§7. Der Rotationsnabel

25. Die vorgelegte Fläche $\mathfrak{x}$ habe jetzt im Nullpunkt einen Nabel. Sollte dieser in einem speziellen Fall in einen Flachpunkt ausgeartet sein, so verwandle man diesen Flachpunkt vermittels einer geeigneten Transformation durch

reziproke Radien in einen gewöhnlichen Nabel. Da bei dieser Transformation bekanntlich Krümmungslinien wieder in Krümmungslinien übergehen, bleibt bei ihr offenbar auch der Index des Nabels invariant. Ist aber der Nullpunkt ein echter Nabel (kein Flachpunkt), so ist die Fläche $\mathfrak{x}$ in der Umgebung dieses Punktes gewiss konvex; die beiden in §5 beschriebenen Abbildungen $\mathfrak{x} \to \mathfrak{X}$ und $\mathfrak{x} \to \mathfrak{U}$ sind mithin umkehrbar eindeutig.

Es sei nun in der Umgebung des Nullpunktes, d.h. etwa im Streifen $0 \leqq \rho \leqq \delta$

$$W(\rho, \theta) = \frac{\rho^2}{4R} + \rho^m(w_0(\theta) + F(\rho, \theta)). \tag{33}$$

Hierbei seien R und m zwei feste positive Zahlen, $m > 2$, $w_0(\theta)$ und $F(\rho, \theta)$ dreimal im ganzen Streifen stetig differentiierbare Funktionen, die in θ die Periode 2π haben. Ausserdem möge $w_0(\theta)$ nicht identisch verschwinden, anderseits sei

$$\lim_{\rho \to 0} F(\rho, \theta) \to 0. \tag{34}$$

Setzt man jetzt die Funktion

$$W_0(\rho, \theta) = \frac{\rho^2}{4R}$$

in die Gleichungen (32) von §6 für W ein, so bemerkt man leicht, dass die so erhaltene Fläche der Gleichung

$$x^2 + y^2 + (z - R)^2 = R^2 \tag{35}$$

genügt und somit eine Kugel vom Radius R ist, die die x,y-Ebene im Nullpunkt berührt.

Andrerseits folgt jetzt aus (33), dass diese Kugel die Fläche $\mathfrak{x}$ im Nullpunkt von der Ordnung $m - 1 > 1$ berührt; mithin ist der Nullpunkt ein Nabel der Fläche und die Kugel (35) die zu diesem Nabel gehörige Schmiegungskugel.[16]

[16] Um die Bedeutung der Voraussetzung, dass sich die Funktion W in der Umgebung des Nabels auf die Form (33) bringen lässt, besser zu verstehen, bemerken wir, dass diese Voraussetzung sicher erfüllt ist, wenn

1. die Ordnung $m - 1 \geqq 2$ der Berührung zwischen Fläche und Schmiegungskugel im Nabel eine ganze Zahl ist,

2. die Funktion W als Funktion von u und v betrachtet im Nullpunkt $m + 1$ mal nach u und v stetig differentiierbar ist.

Dann wird nämlich

$$w_0(\theta) = \sum_{\mu=0}^{m} \frac{\partial^m W}{\partial u^\mu \, \partial v^{m-\mu}}\bigg|_{\substack{u=0 \\ v=0}} \frac{\cos^\mu \theta}{\mu!} \frac{\sin^{m-\mu} \theta}{(m-\mu)!}.$$

Offenbar ist aber die Voraussetzung des Textes allgemeiner, da in der Darstellung (33) für W erstens m eine beliebige positive Zahl > 2, zweitens $w_0(\theta)$ eine beliebige dreimal stetig differentiierbare periodische Funktion sein darf.

26. Setzt man den Ausdruck (33) für die Funktion W in die Formeln (31) für die Koeffizienten A und B der Differentialgleichung (8) der Krümmungslinien ein, so ergibt sich

$$A = \rho^{m-2}((m-1)(w_0' + F_\theta) + \rho F_{\rho\theta}), \tag{36}$$

$$2B = \rho^{m-2}(w_0'' - (m^2 - 2m)(w_0 + F) + F_{\theta\theta} - (2m-1)\rho F_\rho - \rho^2 F_{\rho\rho}), \tag{37}$$

unter w_0', w_0'' die Differentialquotienten von $w_0(\theta)$ verstanden. Ferner folgt aus (34)

$$\lim_{\rho\to 0} F_\theta \to 0, \qquad \lim_{\rho\to 0} F_{\theta\theta} \to 0, \tag{38}$$

$$\lim_{\rho\to 0} \frac{F_\theta}{\rho} \to F_{\rho\theta}(0, \theta). \tag{39}$$

27. I. Fall. Es sei $w_0(\theta)$ gleich einer Konstanten w_0, mithin $w_0' \equiv 0$. Dann lässt sich auf Grund von Formel (36), (37), (38) und (39) die Differentialgleichung (8) durch ρ^{m-1} dividieren, und man erhält nach dieser Division und dem Grenzübergang $\rho \to 0$

$$mF_{\rho\theta}(0, \theta)d\rho^2 - (m^2 - 2m)w_0 d\rho d\theta = 0. \tag{40}$$

Hieraus folgt aber, dass die Differentialgleichung (8) jetzt auch in jedem Punkte der Geraden $\rho = 0$ regulär ist, und zwar ergibt sich aus (40):

$$\left(\frac{d\rho}{d\theta}\right)_1 = 0 \qquad \left(\frac{d\theta}{d\rho}\right)_2 = \frac{F_{\rho\theta}(0, \theta)}{(m-2)w_0}.$$

d.h. aber durch jeden Punkt der Geraden $\rho = 0$ gehen zwei Integralkurven hindurch, deren eine die Gerade $\rho = 0$ selbst ist.

Das bedeutet endlich für die Umgebung des Nabels der vorgelegten Fläche $\mathfrak{x}$, dass, ebenso wie von dem Pol einer Rotationsfläche, von dem Nabel in jeder Richtung θ eine und nur eine Krümmungslinie ausgeht. Diese strahlenförmig von 0 ausgehenden Kurven bilden die eine Schar der Krümmungslinien; die zweite Schar Krümmungslinien auf $\mathfrak{x}$ sind entweder geschlossene Kurven oder Kurven, die sich um den Nullpunkt als asymptotischen Punkt spiralförmig winden. In beiden Fällen ist der Index des Nabels $i = 0$.[17] Damit ist im Falle $w_0 =$ konstant die Behauptung (2) bewiesen.

Man sieht leicht ein, dass die Bedingung $w_0' = 0$ gleichzeitig notwendig und hinreichend dafür ist, dass die Differentialgleichung (8) in jedem Punkt der Geraden $\rho = 0$ regulär ist oder, was dasselbe ist, dass vom Nabel in jeder Richtung eine und nur eine Krümmungslinie ausgeht. Wir nennen einen solchen Nabel einen Rotationsnabel. Er hat die weitere für ihn charakteristische Eigenschaft, dass nur zu einem Rotationsnabel eine Rotationsfläche existiert, welche die vorgelegte Fläche $\mathfrak{x}$ im Nabel von höherer Ordnung, (nämlich der

[17] Vgl. l. c. Fussnote 1; S. 53 Abschnitt **5**.

Ordnung m) berührt als die Schmiegungskugel (35). Diese Rotationsfläche erhält man, wenn man die Funktion

$$W_1 = \frac{\rho^2}{4R} + w_0\rho^m$$

in die Formeln (32) für W einsetzt.[18]

§8. Beweis der Indexabschätzung unter den einfachsten Voraussetzungen

28. II. Fall. Wenn $w_0(\theta)$ nicht gleich einer Konstanten ist, so muss die Ableitung $w_0'(\theta)$ im Intervall $0 \leqq \theta < 2\pi$ mindestens zweimal verschwinden, da in ihm $w_0(\theta)$ als periodische Funktion wenigstens ein Maximum und ein Minimum haben muss.

Wir machen jetzt zunächst die weiteren Voraussetzungen:

(α) *$w_0'(\theta)$ habe im Intervall $0 \leqq \theta < 2\pi$ nur endlich viele Nullstellen $\theta = \theta_\nu$ ($\nu = 1, 2, \cdots, n; n \geqq 2$), die sämtlich einfach sind,*

(β) *der Ausdruck*

$$w_0''(\theta) - (m^2 - 2m)w_0(\theta)$$

sei für alle n Werte $\theta = \theta_\nu$ von Null verschieden.

Jetzt sind, wie man nach Division der Differentialgleichung (8) durch ρ^{m-2} (vgl. die Formeln (36) und (37) für A und B) unmittelbar erkennt, die Stellen $\rho = 0$, $\theta = \theta_\nu$ die einzigen Singularitäten der Differentialgleichung (8), während die Gerade $\rho = 0$ als Doppellösung von (8) durch alle regulären Punkte $\rho = 0$ hindurchgeht, bei denen θ von den n Werten θ_ν verschieden ist. Hieraus schliesst man nach bekannten Sätzen der Theorie der Differentialgleichungen weiter, dass vom Nabel, dem Nullpunkt der Fläche $\mathfrak{x}$, Krümmungslinien nur in den n Richtungen $\theta = \theta_\nu$ ausgehen können. Da wir im folgenden von dieser Tatsache keinen Gebrauch zu machen haben, begnügen wir uns hier mit dieser Andeutung.

29. Aus (36) ergibt sich wegen (38) in Zusammenhang mit Voraussetzung (α)

$$\lim_{\substack{\rho\to 0\\ \theta\to\theta_\nu}} \frac{A_\theta}{\rho^{m-2}} \to (m-1)\, w_0''(\theta_\nu) \neq 0 \tag{41}$$

Aus (41) folgt anderseits, dass sich zu jedem ν für ein hinreichend kleines Intervall $0 \leqq \rho \leqq \delta$ eine Funktion $\theta = \Theta_\nu(\rho)$ eindeutig bestimmen lässt, welche gleichzeitig den beiden Bedingungen

$$\lim_{\rho\to 0} \Theta_\nu(\rho) \to \theta_\nu,\ A(\rho, \Theta_\nu(\rho)) \equiv 0$$

genügt. Mithin existieren im Intervall $0 \leqq \theta < 2\pi$ genau n Kurven $\mathfrak{c}_\nu$, welche die in §1, Voraussetzung 5. beschriebenen Eigenschaften haben und durch die Gleichung $\theta = \Theta_\nu(\rho)$ analytisch dargestellt werden.

---

[18] Vgl. die Arbeit des Verf.: "Zur Theorie der sphärischen Abbildung im Grossen I." *Math. Zeitschr.* **31** (1930), S. 629–708. Siehe insbesondere Fussnote 7; S. 638 und 639. Die zitierte Arbeit behandelt das Problem, eine geschlossene konvexe Fläche mit nur zwei Rotationsnabelpunkten zu konstruieren, wenn das sphärische Bild der Krümmungslinien gegeben ist.

Damit sind aber ausserdem auch die in §4 angegebenen Voraussetzungen 6. und 7. erfüllt, wobei die in Formel (15) auftretenden Exponenten k_ν (eben wegen der Voraussetzung (α)) sämtlich gleich 1 sind. Es kann daher auch keine der Kurven c_ν dritter Art sein, da dann k_ν gerade sein müsste, und zwar ist nach Formel (16), (16') und (17)

$$c_\nu \text{ erster Art, wenn } \operatorname{sg} \frac{A_\theta(\rho, \Theta_\nu(\rho))}{2B(\rho, \Theta_\nu(\rho))} = +1,$$

$$c_\nu \text{ zweiter Art, wenn } \operatorname{sg} \frac{A_\theta(\rho, \Theta_\nu(\rho))}{2B(\rho, \Theta_\nu(\rho))} = -1.$$

Um diese Vorzeichen zu berechnen, berücksichtige man, dass wegen (37), (39) und (41)

$$\operatorname{sg} \frac{A_\theta(\rho, \Theta_\nu(\rho))}{2B(\rho\Theta_\sigma(\rho))} = \operatorname{sg} \lim_{\rho \to 0} \frac{A_\theta(\rho, \Theta_\nu(\rho))}{2B(\rho\Theta_\nu(\rho))}$$

(42)

$$= \operatorname{sg} \frac{(m-1)w_0''(\theta_\nu)}{w_0''(\theta_\nu) - (m^2-2m)w_0(\theta_\nu)} = \operatorname{sg} \frac{1}{1 - (m^2-2m)\dfrac{w_0(\theta_\nu)}{w_0''(\theta_\nu)}},$$

wobei der Nenner wegen Voraussetzung (β) nie gleich Null ist.

30. Nunmehr behaupten wir: Setzt man in den Ausdruck rechter Hand von (42) für θ_ν nach einander die Werte

$$\theta_1 < \theta_2 < \cdots < \theta_\nu < \theta_{\nu+1} < \cdots < \theta_n < \theta_{n+1} = \theta_1 + 2\pi$$

ein, so kann dieser Ausdruck nicht zweimal hintereinander den Wert -1 annehmen, und hieraus folgt dann unmittelbar, wenn wir die Bezeichnung (13) des §3 benutzen,

$$\{+\} \geqq \{-\}, \qquad \text{d.h. aber } i \geqq 0 \text{ wegen (14)}$$

Beweis. Damit

$$\operatorname{sg}\left(1 - (m^2-2m)\frac{w_0(\theta_\nu)}{w_0''(\theta_\nu)}\right) = -1,$$

ist notwendig (aber nicht hinreichend), dass $w_0(\theta_\nu)/w_0''(\theta_2) > 0$. $\operatorname{sg} w_0(\theta_\nu) = \operatorname{sg} w_0''(\theta_\nu)$ bedeutet aber in Verbindung mit $w_0'(\theta_\nu) = 0$, dass die periodische Funktion $w_0(\theta)$ an der Stelle $\theta = \theta_\nu$ entweder ein negatives Maximum oder ein positives Minimum hat. Wenn jetzt gleichzeitig

$$\frac{w_0(\theta_\nu)}{w_0''(\theta_\nu)} > 0, \qquad \frac{w_0(\theta_{\nu+1})}{w_0''(\theta_{\nu+1})} > 0$$

wäre, so würde das heissen, dass etwa auf ein negatives Maximum als nächstes Extremum ein positives Minimum folgen würde (bzw. umgekehrt ein negatives

Maximum auf ein positives Minimum). Da dies unmöglich ist, so ist damit die Behauptung zu Beginn dieses Abschnittes bewiesen.

31. Die im vorigen Abschnitt bewiesene Behauptung lässt sich auch so formulieren, dass zwei auf einander folgende Kurven c_ν, $c_{\nu+1}$ nicht beide zweiter Art sein können, d.h. aber nach den Definitionen der Abschnitte **15.**, **16.**, **17.** §3, dass unter den Voraussetzungen (α) und (β) alle durch die Kurven c_ν, $c_{\nu+1}$ begrenzten Bereiche $\mathfrak{B}_\nu$ von 1. Art oder 3. Art, aber nie von 2. Art sind. Und hieraus folgt nun weiter leicht, dass sich jetzt immer ein aus Integralkurvenstücken zusammengesetzter Polygonzug $\mathfrak{p}$ konstruieren lässt, welcher nur ausspringende Ecken, aber keine einspringenden Ecken mehr enthält. Man braucht zu diesem

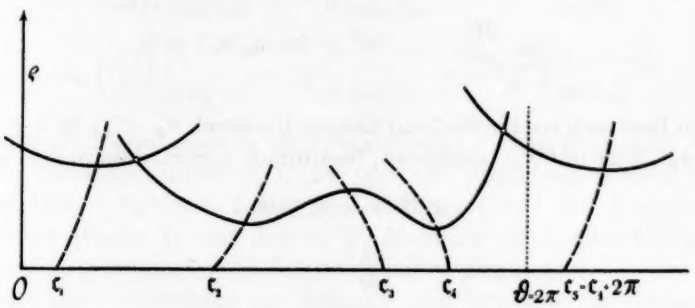

FIG. 5a

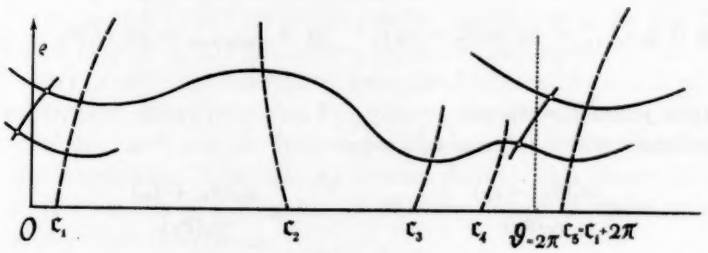

FIG. 5b

Zwecke nur das in Abschnitt **18.** §3 angegebene Konstruktionsverfahren so abzuändern, dass sich der Polygonzug $\mathfrak{p}$ einzig aus den dort mit $\mathfrak{k}_\nu$ bezeichneten Bogen zusammensetzt, was sich unschwer erreichen lässt (vgl. Figur 5a).

Ausgenommen ist hierbei allerdings der Fall, dass sich die eine Schar der Krümmungslinien um den Nabel als asymptotischen Punkt spiralförmig windet. Dann wird sich ein Polygonzug $\mathfrak{p}$ konstruieren lassen, der genau eine einspringende und eine ausspringende Ecke enthält. (vgl. Figur 5b).

§9. Verallgemeinerung der Voraussetzung des §8

32. Wir behalten die Voraussetzung (β) des §8 zunächst noch bei, ersetzen aber die Voraussetzung (α) durch die weniger fordernden, bereits in der Einleitung Abschnitt **7** angegebenen beiden Voraussetzungen

I. $w_0'(\theta)$ *habe nur isolierte Nullstellen,*

II. $w_0(\theta)$ *habe nur Nullstellen erster oder zweiter, aber keine Nullstellen dritter oder höherer Ordnung*; d.h. also, es soll nicht gleichzeitig für ein $\theta = \theta_\nu$

$$w_0(\theta_\nu) = 0, \qquad w_0'(\theta_\nu) = 0 \qquad w_0''(\theta_\nu) = 0$$

sein.

33. Es sei $\theta = \theta_K$ eine Nullstelle höherer Ordnung von $w_0'(\theta)$ d.h. es sei

$$w_0'(\theta_K) = w_0''(\theta_K) = 0.$$

Dann folgt aus Voraussetzung II: $w_0(\theta_K) \neq 0$, und es ist wegen (37)

$$\lim_{\substack{\rho \to 0 \\ \theta \to \theta_K}} \frac{2B}{\rho^{m-2}} \to -(m^2 - 2m) w_0(\theta_K) \neq 0. \tag{43}$$

Ausserdem lässt sich ein hinreichend kleines Rechteck $\theta_K - \epsilon_K \leqq \theta \leqq \theta_K + \epsilon_K$, $0 \leqq \rho \leqq \delta_K$, kurz mit $\mathfrak{R}_K$ bezeichnet, bestimmen derart, dass in ihm wegen (43)

$$\operatorname{sg} B = -\operatorname{sg} w_0(\theta_K) \tag{44}$$

$$\left|\frac{A}{B}\right| < 1 \tag{45}$$

und endlich auf den beiden vertikalen Rechteckseiten wegen (36) und (38)

$$\operatorname{sg} A\,|_{\theta=\theta_K-\epsilon_K} = \operatorname{sg} w_0'(\theta_K - \epsilon_K); \qquad \operatorname{sg} A\,|_{\theta=\theta_K+\epsilon_K} = \operatorname{sg} w_0'(\theta_K + \epsilon_K) \tag{46}$$

ist.

Wir sagen jetzt, $\theta = \theta_K$ sei:

eine Nullstelle erster Art, von w_0', *wenn*

$$-\operatorname{sg} \frac{w_0'(\theta_K - \epsilon_K)}{w_0(\theta_K)} = -1; \qquad -\operatorname{sg} \frac{w_0'(\theta_K + \epsilon_K)}{w_0(\theta_K)} = +1, \tag{47}$$

eine Nullstelle zweiter Art von w_0', *wenn*

$$-\operatorname{sg} \frac{w_0'(\theta_K - \epsilon_K)}{w_0(\theta_K)} = +1; \qquad -\operatorname{sg} \frac{w_0'(\theta_K + \epsilon_K)}{w_0(\theta_K)} = -1 \tag{48}$$

eine Nullstelle dritter Art von w_0', *wenn*

$$-\operatorname{sg} \frac{w_0'(\theta_K - \epsilon_K)}{w_0(\theta_K)} = -\operatorname{sg} \frac{w_0'(\theta_K + \epsilon_K)}{w_0(\theta_K)} \tag{49}$$

Das bedeutet aber: die Nullstelle $\theta = \theta_K$ von w_0' ist:

von erster Art, wenn w_0 *an der Stelle* $\theta = \theta_K$ *ein positives Maximum oder ein negatives Minimum hat,*

von zweiter Art, wenn w_0 *an der Stelle* $\theta = \theta_K$ *ein negatives Maximum oder ein positives Minimum hat,*

von dritter Art, wenn w_0 *an der Stelle* $\theta = \theta_K$ *einen Wendepunkt hat.*

34. Aus (45) folgt jetzt, dass die Reihe (10) des §2 für $(d\rho/d\theta)_1$ im Rechteck $\mathfrak{R}_K$ gleichmässig konvergent ist. Ausserdem kann die Beziehung (10) als eine überall in $\mathfrak{R}_K$ reguläre Differentialgleichung erster Ordnung aufgefasst werden. Mithin lässt sich ein Integralbogen $\mathfrak{k}_K$ konstruieren, der der Differentialgleichung (10) und damit auch (8) genügt und ausserdem zwei Punkte P'_K, P''_K der beiden vertikalen Seiten des Rechteckes $\theta = \theta_K - \epsilon_K$, $\theta = \theta_K + \epsilon_K$ mit einander verbindet, ohne einen Punkt des Geradenstückes $\rho = 0$ zu enthalten, das selbst eine Integralkurve von (10) ist.

Da wegen (12) §2, (44) und (46) für $0 < \rho \leqq \delta_K$

$$\operatorname{sg}\left(\frac{d\rho}{d\theta}\right)_1\Big|_{\theta=\theta_K-\epsilon_K} = \operatorname{sg}\frac{A}{B}\Big|_{\theta=\theta_K-\epsilon_K} = -\operatorname{sg}\frac{w_0'(\theta_K-\epsilon_K)}{w_0(\theta_K)},$$

$$\operatorname{sg}\left(\frac{d\rho}{d\theta}\right)_1\Big|_{\theta=\theta_K+\epsilon_K} = \operatorname{sg}\frac{A}{B}\Big|_{\theta=\theta_K+\epsilon_K} = -\operatorname{sg}\frac{w_0'(\theta_K+\epsilon_K)}{w_0(\theta_K)}$$

ist, so hat der Integralbogen $\mathfrak{k}_K$ im Anfangspunkt P'_K und Endpunkt P''_K Tangenten, deren Vorzeichen durch (47), (48) oder (49) angegeben werden, je nachdem die Nullstelle $\theta = \theta_K$ von 1. Art, 2. Art oder 3. Art ist. Wenn wir nun die eben konstruierten Bogen $\mathfrak{k}_K$ mit den in §3 Abschnitt **18** beschriebenen Bogen $\mathfrak{k}_\nu$ vergleichen, so gelangen wir leicht zu folgendem Ergebnis:

Bezüglich des Vorzeichens der Anfangs- und Endtangenten in den Punkten P'_K und P''_K verhalten sich die Bogen $\mathfrak{k}_K$ genau so, als ob sie

eine Kurve $\mathfrak{c}_K$ von 1. Art schneiden, wenn die Nullstelle θ_K von 1. Art ist,
eine Kurve $\mathfrak{c}_K$ von 2. Art schneiden, wenn die Nullstelle θ_K von 2. Art ist,
eine Kurve $\mathfrak{c}_K$ von 3. Art schneiden, wenn die Nullstelle θ_K von 3. Art ist.

Man konstruiere daher jetzt den Polygonzug $\mathfrak{p}$ genau ebenso wie in Abschnitt **18.** §3, wobei man sich nur die Rechtecke $\mathfrak{R}_K$ durch eine Kurve $\mathfrak{c}_K$ von gleicher Art wie die zugehörige Nullstelle θ_K ersetzt denkt. Bei dieser Konstruktion sind wieder die Kurven $\mathfrak{c}_K$ 3. Art ohne Bedeutung, sodass man sie sich ebenso wie in §3 forgelassen denken kann.

Nachdem man so die Hilfskurven $\mathfrak{c}_K$, soweit sie 1. oder 2. Art sind, zu den den einfachen Nullstellen $\theta = \theta_\nu$ von w_0 entsprechenden Kurven $\mathfrak{c}_\nu$ hinzugefügt hat, zeigt man ebenso wie in Abschnitt **30.** §8, dass zwei Kurven $\mathfrak{c}_\nu$ 2. Art nicht unmittelbar aufeinander folgen können; und damit ist die Behauptung $i \geqq 0$ unter den Voraussetzungen (β) (Abschnitt **28.** §8), I und II (Abschnitt **31.** §9) bewiesen.

35. Wir lassen jetzt die Voraussetzung (β) fallen, nehmen somit an, es gäbe eine Nullstelle $\theta = \theta_\mu$ von w_0', bei der

$$w_0''(\theta_\mu) - (m^2 - 2m)w_0(\theta_\mu) = 0. \tag{50}$$

Aus Voraussetzung II (Abschnitt **31.**) folgt dann

$$w_0''(\theta_\mu) \neq 0 \qquad w_0(\theta_\mu) \neq 0$$

und weiter wegen (50)

$$\frac{w_0(\theta_\mu)}{w_0''(\theta_\mu)} > 0. \tag{51}$$

Wegen (41) existiert nach dem Satz über implizite Funktionen eine eindeutig bestimmte Funktion $\theta = \Theta_\mu(\rho)$, die gleichzeitigden beiden Bedingungen

$$A(\rho, \Theta_\mu(\rho)) = 0, \qquad \lim_{\rho=0} \Theta_\mu \to \theta_\mu$$

genügt. Die durch diese Funktion dargestellte Kurve $\mathfrak{c}_\mu$ ist nun 1. Art oder 2. Art, je nachdem

$$\operatorname{sg} \frac{A_\theta}{2B}\Big|_{\theta=\Theta_\mu(\rho)} = +1 \text{ oder } \operatorname{sg} \frac{A_\theta}{2B}\Big|_{\theta=\Theta_\mu(\rho)} = -1$$

ist. In beiden Fällen folgt aber aus (51), dass die Stelle $\theta = \theta_\mu$ entweder ein positives Minimum oder ein negatives Maximum der Funktion $w_0(\theta)$ ist, d.h. aber, dass weder die unmittelbar vorausgehende Kurve $\mathfrak{c}_{\mu-1}$ noch die unmittelbar folgende Kurve $\mathfrak{c}_{\mu+1}$ von 2. Art sein kann. Damit ist die Behauptung $i \geqq 0$ unter den Voraussetzungen I und II vollständig bewiesen.

36. Aus der soeben bewiesenen Tatsache: unter den Voraussetzungen I und II können nie zwei Kurven 2. Art $\mathfrak{c}_\nu$ und $\mathfrak{c}_{\nu+1}$ unmittelbar auf einander folgen, lässt sich ebenso wie in Abschnitt **31.** §8 leicht schliessen, *dass sich um den Nabel immer ein aus Stücken von Krümmungslinien zusammengesetzter geschlossener Polygonzug* $\mathfrak{p}^*$ *konstruieren lässt, welcher überhaupt keine einspringenden Ecken enthält; ausgenommen ist hierbei wieder nur der Fall, dass sich die eine Schar der Krümmungslinien um den Nabel als asymptotischen Punkt spiralförmig windet.*

In dem später zu veröffentlichenden II. Teil wird die einschränkende Voraussetzung II fallen gelassen; es sind dann somit Funktionen $w_0(\theta)$ zugelassen, die Nullstellen beliebig hoher Ordnung haben. Dann wird man aber zeigen können, dass es Fälle gibt, in denen zwei unmittelbar aufeinander folgende Kurven $\mathfrak{c}_\nu$ zweiter Art auftreten, die beide vom gleichen Punkte $\theta = \theta_\nu$ der Geraden $\rho = 0$ ausgehen. In diesem Falle lässt sich um den Nabel nicht ohne weiteres ein Polygonzug $\mathfrak{p}^*$ konstruieren der nur ausspringende Ecken enthält. Die Tatsache, dass dieser Ausnahmefall auftreten kann, ist wohl der Hauptgrund dafür, dass die Untersuchungen des II. Teils wesentlich schwieriger und langwieriger sind als im I. Teil.

BERLIN, GERMANY.

ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

ON THE TRANSITIVITY OF PERSPECTIVE MAPPINGS

BY J. v. NEUMANN AND I. HALPERIN[1]

(Received July 8, 1939)

A system L of elements $a, b, c, x, y, \cdots$ is said to be *partially ordered* if a relation $a \leqq b$ (equivalently, $b \geqq a$) is defined for certain pairs of elements in such a way that

I_1: $a \leqq b, b \leqq a$ *if and only if* $a = b$.

I_2: $a \leqq b, b \leqq c$ *together imply* $a \leqq c$.

A subset S of L is said to have a sum (intersection) in L, written $\sum (S)(\prod (S))$, if for every c in L, $c \geqq \sum (S)$ $(c \leqq \prod (S))$ if and only if $c \geqq a (c \leqq a)$ for every a in S.[2]

A partially ordered L is called a *lattice* if

II: *Every finite subset* $a_1, a_2, \cdots, a_n$ *has a sum* $\sum_{i=1}^{n} a_i$ *(written also* $a_1 + a_2 + \cdots + a_n$*) and an intersection* $\prod_{i=1}^{n} a_i$ *(written also* $a_1 a_2 \cdots a_n$*).*[3]

A lattice is said to be *countably continuous* if

III_1': *Every countable subset* $a_1, a_2, \cdots$ *has a sum* $\sum_{i=1}^{\infty} a_i$ *and for every* c *in* L, $c \sum_{i=1}^{\infty} a_i = \sum_{n=1}^{\infty} (c \sum_{i=1}^{n} a_i)$.

III_2': *Every countable subset* $a_1, a_2, \cdots$ *has an intersection* $\prod_{i=1}^{\infty} a_i$ *and for every* c *in* L, $c + \prod_{i=1}^{\infty} a_i = \prod_{n=1}^{\infty} (c + \prod_{i=1}^{\infty} a_i)$.

A lattice is said to be *continuous* if

III_1: *For every ordinal* Ω *and for every set of elements* a_α, $\alpha < \Omega$, *there is a sum* $\sum_{\alpha<\Omega} a_\alpha$, *and for every* c *in* L, $c \sum_{\alpha<\Omega} a_\alpha = \sum_{\gamma<\Omega} (c \sum_{\alpha\leqq\gamma} a_\alpha)$.[4]

III_2: *For every ordinal* Ω *and for every set of elements* a_α, $\alpha < \Omega$, *there is an intersection* $\prod_{\alpha<\Omega} a_\alpha$, *and for every* c *in* L, $c + \prod_{\alpha<\Omega} a_\alpha = \prod_{\gamma<\Omega} (c + \prod_{\alpha\leqq\gamma} a_\alpha)$.[4]

A lattice is called *modular* if

IV: $a \leqq c$ *implies that* $(a + b)c = a + bc$ *for all* b *in* L.

A lattice is called *complemented if*

V: *Whenever* $a \leqq b \leqq c$ *there exists an element* d *(not necessarily unique) such that* $b + d = c$ *and* $bd = a$. *(Such a* d *will be written* $[c - b]_a$.)

Continuous complemented modular lattices were defined and studied by one of the writers and it was shown that any such lattice which satisfies a certain irreducibility condition is either a projective geometry or a geometry of a new,

[1] Sterling Research Fellow, Yale University.

[2] The sum (intersection) is clearly unique if it exists at all.

[3] This would be implied if every pair of elements a, b had a sum $a + b$ and an intersection ab.

[4] By the well-ordering theorem every set of elements a_t, t ranging over an arbitrary set of subscripts, can be arranged as a set a_α, $\alpha < \Omega$, for some ordinal Ω.

non-finite dimensional type which was named *continuous geometry.*[5] Of great importance in this work are the notion of *perspectivity* (a is *perspective* to b, written $a \sim b$, if for some x, called an *axis*, $a + x = b + x$ and $ax = bx$), and the transitivity of the relation of perspectivity, that is, for any finite n,

(*) $a_1 \sim a_2$, $a_2 \sim a_3$, $\cdots$, $a_n \sim a_{n+1}$ *together imply* $a_1 \sim a_{n+1}$.

The general validity of (*) was established by proving first the special case

(**) $a_1 \sim a_2$, $a_2 \sim a_3$, $\cdots$, $a_n \sim a_{n+1}$, $a_1 a_{n+1} = \prod_{i=1}^{n+1} a_i$, *together imply* $a_1 \sim a_{n+1}$.[6]

In the present note we give simple proofs of theorems concerning perspective mappings in modular lattices, from which (**) follows. While the previous proof of (**) is valid in every countably continuous complemented modular lattice[7] the present proof shows that (**) holds in every complemented modular lattice which satisfies III_1 (but not necessarily III_2).

In what follows we assume only that L is a modular lattice.

DEFINITION 1.[8] *$L(a, b)$ is the set of all c in L for which $a \leqq c \leqq b$.*

LEMMA 1. *If $a \sim b$ with axis x and $c = ax = bx$, then a $(1, 1)$ correspondence between the elements of $L(c, a)$ and those of $L(c, b)$ which preserves the relation $\leqq$ (hence also the operations of sum and intersection) is defined by the inverse mappings*

$$(T) \qquad a_1 \to Ta_1 = (a_1 + x)b \quad \text{for} \quad c \leqq a_1 \leqq a.$$

$$(S) \qquad b_1 \to Sb_1 = (b_1 + x)a \quad \text{for} \quad c \leqq b_1 \leqq b.$$

PROOF. Since $(a_1 + x)b \geqq bx = c$ and $(a_1 + x)b \leqq b$ it follows that Ta_1 is an element of $L(c, b)$. If a_1 is an element of $L(c, a)$ then, from the modularity of L, we have

$$STa_1 = \{(a_1 + x)b + x\}a = (a_1 + x)(b + x)a = (a_1 + x)(a + x)a$$
$$= (a_1 + x)a = a_1 + ax = a_1$$

that is, $STa_1 = a_1$. Similarly, $TSb_1 = b_1$ if b_1 is an element of $L(c, b)$. Thus T, S are inverse mappings which define a (1, 1) correspondence as stated. It is clear from the definition of T, S that the relation $\leqq$ is preserved.

DEFINITION 2. *The mapping T of Lemma 1 is called a "perspective mapping." The axis x is said to determine the mapping.*

---

[5] See J. v. Neumann: *Proceedings of the National Academy of Sciences*, vol. 32 (1936), pp. 92–100, and other references given there. The axioms formulated there are slightly different from but equivalent to those used here.

[6] See J. v. Neumann: *Continuous Geometry*, planographed, Institute for Advanced Study, Princeton, N. J., 1936–38, particularly Part I, Theorem 4.2.

[7] Even the general (*) holds in any countably continuous complemented modular lattice. See I. Halperin: *Transactions of the American Mathematical Society*, vol. 44 (1938), pp. 537–62.

[8] Definitions 1–5, Lemmas 1–5, 8, 10 and Theorem 1 are taken from Part I of the work referred to in footnote (6). The "independence over c" defined in Definition 3, would be "independence" as defined in loc. cit. if c were a "zero" or "least" element in L.

LEMMA 2. *If the axis x determines a perspective mapping T of $L(c, a)$ on $L(c, b)$ then $x(a + b)$ determines the same mapping.*

PROOF. $x(a + b)$ is an axis of perspectivity between a and b since: $a + x(a + b) = (a + x)(a + b) = a + b$; similarly, $b + x(a + b) = a + b = a + x(a + b)$; and $ax(a + b) = ax = c$; similarly, $bx(a + b) = c$. Now $x(a + b)$ determines the mapping T since for all $c \leqq a_1 \leqq a$ we have $\{a_1 + x(a + b)\}b = (a_1 + x)(a + b)b = (a_1 + x)b = Ta_1$.

LEMMA 3. *If T is a perspective mapping of $L(c, a)$ on $L(c, b)$ and $c \leqq a_1 \leqq a$ then the T-mapping of $L(c, a_1)$ on $L(c, Ta_1)$ is a perspective mapping; if $c \leqq a_1 \leqq ab$ then this T-mapping is the identity mapping of $L(c, a_1)$ on itself.*

PROOF. Let x be the axis determining T. Then

$$a_1x = a_1ax = c; \qquad (Ta_1)x = (a_1 + x)bx = c;$$

$$a_1 + x = (a_1 + x)(a + x) = (a_1 + x)(b + x) = (a_1 + x)b + x = (Ta_1) + x;$$

hence $a_1 \sim Ta_1$ with axis x. Now the T-mapping of $L(c, a_1)$ is the perspective mapping determined by the axis x since for all $c \leqq a_2 \leqq a_1$, we have $(a_2 + x)Ta_1 = (a_2 + x)(a_1 + x)b = (a_2 + x)b = Ta_2$. If $c \leqq a_1 \leqq ab$ it is clear that for all $c \leqq a_2 \leqq a_1$ we have $Ta_2 = (a_2 + x)b + a_2 + bx = a_2$.

DEFINITION 3. *The elements a, d, b are said to be "independent"* (over c) if $a(d + b) = d(a + b) = b(a + d)(= c)$.

It is clear that if a, d, b are independent, they are independent over a unique c.

LEMMA 4. *a, d, b are independent (over c) if and only if $ab = d(a + b)(= c)$.*

PROOF. The necessity of the conditions is clear since from Definition 3, $d(a + b) = c$ and $ab = ab(a + d)(b + d) = cc = c$. On the other hand, if the conditions are satisfied, then $a(d + b) = a(a + b)(d + b) = a\{b + d(a + b)\} = a(b + ab) = ab = c$; similarly, $b(a + d) = c$. Hence a, d, b are independent over c.

LEMMA 5. *If a, d, b are independent then $(a + d)(b + d) = d$.*

PROOF. $(a + d)(b + d) = d + a(b + d) = d + d(a + b) = d$.

THEOREM 1. *Let T_1 be a perspective mapping of $L(c, a)$ on $L(c, d)$, let T_2 be a perspective mapping of $L(c, d)$ on $L(c, b)$, and let T be the product mapping T_2T_1. If a, d, b are independent then T is a perspective mapping.*

PROOF. By Lemma 2 there are axes x_1, x_2 determining T_1, T_2 respectively, such that $x_1 \leqq (a + d)$ and $x_2 \leqq (b + d)$. Define $x = x_1 + x_2$. Then $a \sim b$ with axis x since:

$$ax = a(a + d)(x_1 + x_2) = a\{x_1 + (a + d)x_2(b + d)\}$$
$$= a(x_1 + dx_2) = ax_1 = c;$$

similarly, $bx = c$; and $a + x = a + x_1 + x_2 = d + x_1 + x_2 = b + x_1 + x_2 = b + x$. Now T is the perspective mapping of $L(c, a)$ on $L(c, b)$ determined by the axis x since for all $c \leqq a_1 \leqq a$ we have

$$(a_1 + x)b = (a_1 + x_1 + x_2)(b + d)b = \{(a_1 + x_1)(a + d)(b + d) + x_2\}b$$
$$= \{(a_1 + x_1)d + x_2\}b = T_2T_1a_1 = Ta_1.$$

LEMMA 6. *Let T be a perspective mapping of $L(c, a)$ on $L(c, b)$ and let T_1 be a perspective mapping of $L(c, d)$ on $L(c, a)$. If $d(a + b) \leqq ab$ then there exists a perspective mapping T_2 of $L(c, d)$ on $L(c, b)$ such that $T = T_2(T_1)^{-1}$.*

PROOF. Let x be an axis determining T such that $x \leqq (a + b)$; let x_1 be an axis determining T_1 such that $x_1 \leqq (a + d)$ (x, x_1 exist by Lemma 2). Define $x_2 = x + x_1$. Then x_2 determines a perspective mapping T_2 of $L(c, d)$ on $L(c, b)$ since:

$$d + x_2 = d + x + x_1 = a + x + x_1 = b + x + x_1 = b + x_2; \text{ and}$$

$$dx_2 = d(a + d)(x + x_1) = d\{x_1 + (a + b)x(a + d)\}$$

$$= d\{x_1 + x(a + d(a + b))\} = d(x_1 + ax) = dx_1 = c;$$

similarly,

$$bx_2 = b(b + a)(x + x_1) = b\{x + (a + b)x_1(a + d)\}$$

$$= b(x + ax_1) = bx = c.$$

Now $T = T_2(T_1)^{-1}$ since for all $c \leqq a_1 \leqq a$ we have

$$\{(a_1 + x_1)d + x_2\}b = \{(a_1 + x_1)d + x + x_1\}b = \{(a_1 + x_1)(x_1 + d) + x\}b$$

$$= \{(a_1 + x_1)(a + d) + x\}b = (a_1 + x_1 + x)(a + b)b$$

$$= \{(a_1 + x) + (a + d)x_1(a + b)\}b = (a_1 + x + ax_1)b$$

$$= (a_1 + x)b = Ta_1.$$

THEOREM 2. *Let T_i be a perspective mapping of $L(c, a_i)$ on $L(c, a_{i+1})$ for $i = 1, \cdots, n$. If there exists an element d with $d(a_i + a_{i+1}) \leqq a_i a_{i+1}$ for $i = 1, \cdots, n$ and $d(a_1 + a_{n+1}) = a_1 a_{n+1}$, and a perspective mapping T_1' of $L(c, d)$ on $L(c, a_1)$, then the product mapping $T = T_n T_{n-1} \cdots T_2 T_1$ is a perspective mapping of $L(c, a_1)$ on $L(c, a_{n+1})$.*

PROOF. By successive applications of Lemma 6 we can find perspective mappings T_i' of $L(c, d)$ on $L(c, a_i)$ such that $T_i = T_{i+1}'(T_i')^{-1}$. Then

$$T = T_n T_{n-1} \cdots T_2 T_1$$

$$= T_{n+1}'(T_n')^{-1} \cdot T_n'(T_{n-1}')^{-1} \cdots T_3'(T_2')^{-1} T_2'(T_1')^{-1} = T_{n+1}'(T_1')^{-1}.$$

Since $d(a_1 + a_{n+1}) = a_1 a_{n+1}$ Lemma 4 shows that a_1, d, a_{n+1} are independent. Theorem 1 then shows that T is a perspective mapping as stated.

THEOREM 3. *Let T_1 be a perspective mapping of $L(c, a)$ on $L(c, d)$, let T_2 be a perspective mapping of $L(c, d)$ on $L(c, b)$, and let T be the product mapping $T_2 T_1$. If $ad = db = c$ and $a_1 + T_1 a_1 = Ta_1 + T_1 a_1$ for all $c \leqq a_1 \leqq a$ then T is a perspective mapping.*

PROOF. Since $T_1 a = d$ and $Ta = b$ we have $a + d = d + b$ and $ad = db = c$; hence d determines a perspective mapping of $L(c, a)$ on $L(c, b)$. This mapping is identical with T since for all $c \leqq a_1 \leqq a$ we have $(a_1 + d)b = (a_1 + T_1 a_1 + d)b = (Ta_1 + T_1 a_1 + d)b = Ta_1 + db = Ta_1$.

DEFINITION 4. *$a < b$ means that $a \leqq b$ but $a \neq b$.*

LEMMA 7. *Let T_1 be a perspective mapping of $L(c, a)$ on $L(c, d)$, let T_2 be a perspective mapping of $L(c, d)$ on $L(c, b)$ and let T be the product mapping T_2T_1. If $c < a$ and L is complemented then there exists an a_1 with $c < a_1 \leqq a$ such that that the T-mapping of $L(c, a_1)$ on $L(c, Ta_1)$ is a perspective mapping.*

PROOF. If $c < ad$ we can take $a_1 = ad$; for Lemma 3 shows that the T_1-mapping of $L(c, ad)$ is the identity mapping and that the T_2-mapping, which then coincides with the T-mapping, is a perspective mapping. Since $c \leqq ad$ we need therefore consider only the case $c = ad$. Similarly, we may suppose that $c = db$. If now $a_2 + T_1a_2 = Ta_2 + T_1a_2$ for all $c \leqq a_2 \leqq a$ then Theorem 3 shows that T itself is a perspective mapping. If for some $c \leqq a_2 \leqq a$ we have $a_2 + T_1a_2 \neq Ta_2 + T_1a_2$ then either (i) $a_2(Ta_2 + T_1a_2) < a_2$ or (ii) $(Ta_2)(a_2 + T_1a_2) < Ta_2$. If (i) holds, define $a_1 = [a_2 - a_2(Ta_2 + T_1a_2)]_c$. Then $c < a_1 \leqq a_2$, hence $(Ta_1)(T_1a_1) = (Ta_1)b(T_1a_1)d = c$ and $a_1(Ta_1 + T_1a_1) = a_1(Ta_1 + T_1a_1)(Ta_2 + T_1a_2) = c$. Then Lemma 4 shows that a_1, T_1a_1, Ta_1 are independent. Now Lemma 3 and Theorem 1 show that the T-mapping of $L(c, a_1)$ is a perspective mapping. If (ii) holds, define $a_1 = T^{-1}[(Ta_2) - (Ta_2)(a_2 + T_1a_2)]_c$. Then $c < a_1 \leqq a_2$, $a_1(T_1a_1) = c$, and $(Ta_1)(a_1 + T_1a_1) = (Ta_1)(Ta_2)(a_1 + T_1a_1)(a_2 + T_1a_2) = c$. Hence a_1, T_1a_1, Ta_1 are independent and the T-mapping of $L(c, a_1)$ on $L(c, Ta_1)$ is a perspective mapping. This proves the lemma.

LEMMA 8. *If T_i is a perspective mapping of $L(c, a_i)$ on $L(c, a_{i+1})$ for $i = 1, \cdots, n$ and T is the product mapping $T_nT_{n-1} \cdots T_2T_1$, then T is a (1, 1) mapping of $L(c, a_1)$ on $L(c, a_{n+1})$ which preserves the relation $\leqq$.*

PROOF. This is immediate from Definition 2.

DEFINITION 5. *The mapping T of Lemma 8 is called a "projective mapping."*

LEMMA 9. *Let T be a projective mapping of $L(c, a)$ on $L(c, b)$. If $c < a$ and L is complemented then there exists an element a_1 with $c < a_1 \leqq a$ such that the T-mapping of $L(c, a_1)$ on $L(c, Ta_1)$ is a perspective mapping.*

PROOF. This follows from successive applications of Lemma 7.

DEFINITION 6. *A (1, 1) mapping T of $L(c, a)$ on $L(c, b)$ which preserves the relation $\leqq$, is called "piece-wise perspective" if for some ordinal Ω a can be expressed as a sum $a = \sum_{\alpha<\Omega} u_\alpha$, where for every $\alpha < \Omega$*

(i)$_\alpha$ *$c \leqq u_\alpha \leqq a$ and, if $\alpha > 1$, $c < u_\alpha$.*

(ii)$_\alpha$ *$u_\alpha u^\alpha = c$ for some u^α with $a \geqq u^\alpha \geqq u_\beta$ for all $\beta < \alpha$.*

(iii)$_\alpha$ *The T-mapping of $L(c, u_\alpha)$ on $L(c, Tu_\alpha)$ is a perspective mapping.*

THEOREM 4. *Every projective mapping in a complemented modular lattice is piece-wise perspective.*

PROOF. Let T be a projective mapping of $L(c, a)$ on $L(c, b)$. Now define $u_1 = u^1 = c$ and define u_α, u^α for $\alpha > 1$ by transfinite induction as follows: Suppose that for some ordinal γ we have already defined u_α, u^α to satisfy (i)$_\alpha$, (ii)$_\alpha$, (iii)$_\alpha$ of Definition 6 for all $\alpha < \gamma$. If there exists an element d such that $d \geqq u_\alpha$ for all $\alpha < \gamma$ and $d \geqq a$ does not hold, then define $u^\gamma = ad$, and $y = [a - u^\gamma]_c$. Then $c < y$. From Lemmas 3, 9 there exists an element u_γ such

that $c < u_\gamma \leqq y$ and the T-mapping of $L(c, u_\gamma)$ is a perspective mapping. It is clear that $u_\gamma u^\gamma = u_\gamma u^\gamma y = c$ and hence $(\mathrm{i})_\gamma$, $(\mathrm{ii})_\gamma$, $(\mathrm{iii})_\gamma$ are satisfied. But from $(\mathrm{i})_\alpha$, $(\mathrm{ii})_\alpha$, it follows that the u_α are mutually different and hence cannot have greater cardinal power than the set of all elements in L. Thus it must happen for some Ω that no $d \geqq u_\alpha$ for all $\alpha < \Omega$ exists such that $d \geqq a$ does not hold, that is, $a = \sum_{\alpha<\Omega} u_\alpha$. This proves the theorem.

LEMMA 10. *If $a \sim b$ and c is an element $\leqq ab$ and L is complemented, then there exists a perspective mapping of $L(c, a)$ on $L(c, b)$.*

PROOF. Let $a \sim b$ with axis x and define $y = [(x + ab) - ab]_c$. Then $y \leqq (x + ab)$, hence $ay = a(x + ab)y = (ax + ab)y = aby = c$; similarly, $by = c$; and $a + y = a + ab + y = a + x + ab = a + x$; similarly, $b + y = b + x = a + x$. The lemma now follows from Lemma 1.

THEOREM 5. *Let $a_1 \sim a_2$, $a_2 \sim a_3$, $\cdots$, $a_n \sim a_{n+1}$ hold in a complemented modular lattice. If there exists an element d such that $d \sim a_1$, $d(a_i + a_{i+1}) \leqq a_i a_{i+1}$ for all $i = 1, \cdots, n$ and $d(a_1 + a_{n+1}) = a_1 a_{n+1}$, then $a_1 \sim a_{n+1}$.*

PROOF. Let $c = a_1 a_{n+1}$. Then clearly $c \leqq d$ and $c \leqq a_1$. If $c \leqq a_i$ for any i it follows that $a_{i+1} \geqq a_i a_{i+1} \geqq d(a_i + a_{i+1}) \geqq cc = c$, that is $c \leqq a_{i+1}$. Thus by induction on i we have $c \leqq a_i$ for all $i = 1, \cdots, n + 1$. From Lemma 10 it follows that there exists a perspective mapping of $L(c, a_i)$ on $L(c, a_{i+1})$ for all $i = 1, \cdots, n$ and of $L(c, d)$ on $L(c, a_i)$. The theorem now follows from Theorem 2.[9]

THEOREM 6. *Let $a_1 \sim a_2$, $a_2 \sim a_3$, $\cdots$, $a_n \sim a_{n+1}$ hold in a complemented modular lattice which satisfies III_1.*[10] *If $a_1 a_{n+1} = \prod_{i=1}^{n+1} a_i$ then $a_1 \sim a_{n+1}$.*

PROOF. Define $c = a_1 a_{n+1}$. Then by Lemma 10 there exists a perspective mapping of $L(c, a_i)$ on $L(c, a_{i+1})$ for all $i = 1, \cdots, n$, hence a projective mapping T of $L(c, a_1)$ on $L(c, a_{n+1})$. It follows from Theorem 4 that we can express a_1 as $a_1 = \sum_{\alpha<\Omega} u_\alpha$ where $c \leqq u_\alpha \leqq a_1$, $u_\alpha(\sum_{\beta<\alpha} u_\beta) = u_\alpha u^\alpha(\sum_{\beta<\alpha} u_\beta) = c$, and the T-mapping of $L(c, u_\alpha)$ is a perspective mapping for all $\alpha < \Omega$. By Lemma 2 we may assume that the T-mapping of $L(c, u_\alpha)$ is determined by an x_α such that $x_\alpha \leqq (u_\alpha + Tu_\alpha)$. Now define $x = \sum_{\alpha<\Omega} x_\alpha$. Since

$$(Tu_\gamma)(\textstyle\sum_{\alpha<\gamma} Tu_\alpha) = T(u_\gamma \sum_{\alpha<\gamma} u_\alpha) = Tc = c,$$

and $(\sum_{\alpha\leqq\gamma} u_\alpha)(Tu_\gamma + \sum_{\alpha<\gamma} Tu_\alpha) = (\sum_{\alpha\leqq\gamma} u_\alpha)(Tu_\gamma + \sum_{\alpha<\gamma} Tu_\alpha)a_1 a_{n+1} = c$, it follows by Lemma 4 that $(\sum_{\alpha\leqq\gamma} u_\alpha)(\sum_{\alpha<\gamma} Tu_\alpha)$, Tu_γ are independent over c. Similarly, $(\sum_{\alpha<\gamma} u_\alpha)$, u_γ, $(\sum_{\alpha<\gamma} Tu_\alpha)$ are independent over c. Then

$$\{(\textstyle\sum_{\alpha<\gamma} x_\alpha) + x_\gamma\}\{(\sum_{\alpha<\gamma} u_\alpha) + u_\gamma\}$$
$$= \{(\textstyle\sum_{\alpha<\gamma} x_\alpha) + x_\gamma(\sum_{\alpha<\gamma} x_\alpha + \sum_{\alpha<\gamma} u_\alpha + u_\gamma)\}\{(\textstyle\sum_{\alpha<\gamma} u_\alpha) + u_\gamma\}$$

[9] This shows that (**) holds in any modular lattice which can be suitably enlarged to provide an element d with $d \sim a_1$ and $d(\sum_{i=1}^{n+1} a_i) = a_1 a_{n+1}$.

[10] In such a lattice there exists a zero element and an extensive theory of independence (see loc. cit. footnote (6)). The use of this theory would simplify the proof of Theorem 6.

$$= \{(\sum_{\alpha<\gamma} x_\alpha) + x_\gamma(u_\gamma + Tu_\gamma)(u_\gamma + \sum_{\alpha<\gamma} u_\alpha + \sum_{\alpha<\gamma} Tu_\alpha)\}(\sum_{\alpha\leqq\gamma} u_\alpha)$$
$$= \{(\sum_{\alpha<\gamma} x_\alpha) + x_\gamma[u_\gamma + (Tu_\gamma)(\sum_{\alpha\leqq\gamma} u_\alpha + \sum_{\alpha<\gamma} Tu_\alpha)]\}(\sum_{\alpha\leqq\gamma} u_\alpha)$$
$$= (\sum_{\alpha<\gamma} x_\alpha)(\sum_{\alpha\leqq\gamma} u_\alpha)$$
$$= (\sum_{\alpha<\gamma} x_\alpha)(u_\gamma + \sum_{\alpha<\gamma} u_\alpha)(\sum_{\alpha<\gamma} u_\alpha + \sum_{\alpha<\gamma} Tu_\alpha)$$
$$= (\sum_{\alpha<\gamma} x_\alpha)(\sum_{\alpha<\gamma} u_\alpha).$$

That is, $(\sum_{\alpha\leqq\gamma} x_\alpha)(\sum_{\alpha\leqq\gamma} u_\alpha) = (\sum_{\alpha<\gamma} x_\alpha)(\sum_{\alpha<\gamma} u_\alpha)$.
Then, using III_1, we have

$$\begin{aligned}(\sum_{\alpha<\beta} x_\alpha)(\sum_{\alpha<\beta} u_\alpha) &= \sum_{\gamma_1<\beta}(\sum_{\alpha\leqq\gamma_1} x_\alpha)(\sum_{\alpha<\beta} u_\alpha)\\ &= \sum_{\gamma_1<\beta}\sum_{\gamma_2<\beta}(\sum_{\alpha\leqq\gamma_1} x_\alpha)(\sum_{\alpha\leqq\gamma_2} u_\alpha)\\ &= \sum_{\gamma<\beta}(\sum_{\alpha\leqq\gamma} x_\alpha)(\sum_{\alpha\leqq\gamma} u_\alpha)\\ &= \sum_{\gamma<\beta}(\sum_{\alpha<\gamma} x_\alpha)(\sum_{\alpha<\gamma} u_\alpha).\end{aligned}$$

Since $x_1u_1 = c$ it follows by transfinite induction on β that $(\sum_{\alpha<\beta} x_\alpha)(\sum_{\alpha<\beta} u_\alpha) = c$ for all β. Hence, setting $\beta = \Omega$, we have $xa_1 = c$. Similarly, $xa_{n+1} = c$. Finally,

$$a_1 + x = (\sum_{\alpha<\Omega} u_\alpha) + (\sum_{\alpha<\Omega} x_\alpha) = (\sum_{\alpha<\Omega} Tu_\alpha) + (\sum_{\alpha<\Omega} x_\alpha) = a_{n+1} + x.$$

Thus $a_1 \sim a_{n+1}$ with axis x and the theorem is established.

COROLLARY. *Let $a_1 \sim a_2$, $a_2 \sim a_3$, $\cdots$, $a_n \sim a_{n+1}$ hold in a complemented modular lattice which satisfies III_1. If $a_1a_{n+1} = 0$,[11] then $a_1 \sim a_{n+1}$.*

PROOF. This is clearly a particular case of Theorem 6.

THE INSTITUTE FOR ADVANCED STUDY, AND
HARVARD UNIVERSITY

[11] 0 denotes the zero element of L (see footnote (10)).

ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

ON RINGS OF OPERATORS. III

BY J. v. NEUMANN

(Received June 14, 1939)

INTRODUCTION

1. In two earlier papers ((**1**), (**2**)), F. J. Murray and the author investigated certain operator rings, called *factors*. (Cf. (**1**), p. 138, Definition 3.1.2. For an introduction into the theory of operator rings cf. (**5**) particularly pp. 372–376, 388–398.) The motives which led to those investigations are described in (**1**), pp. 116–123.

The main principle of classification for factors, which was found in (**1**), is based on the ranges of their *relative dimension functions*. (Cf. *ibid.*, p. 165, Definition 8.2.1; p. 168, Theorem VII; and p. 172, Theorem VIII.) Thus all factors were found to belong to classes called (I_n) $(n = 1, 2, \cdots)$, (I_∞), (II_1), (II_∞), (III_∞). It was shown that factors in each one of these classes actually do exist—with the exception of (III_∞), which had to be left undecided. (Cf. (**1**), p. 208, Theorem XII.)

The main result of this paper is that factors of class (III_∞) also exist. (Cf. Theorem IX, and Examples (α), (β) in §4.4.) Thus the above classification of factors is fully justified.

2. This result is obtained by the simultaneous use of two devices: The notions of *norm* and *normedness*, which form the subject of Chapter I; and the process $A^{|E_1|E_2|\cdots}$ which is the analogue of the process of taking the diagonal part of a finite matrix, and is discussed in Chapter II.

Both notions have an interest of their own, and we think that they deserve to be studied for their own sake. In this paper their analysis is restricted to the amount which is necessary for our immediate purposes, although we permitted ourselves a certain leeway in some cases where a little more generality than absolutely necessary seemed to make things clearer and easier. But we propose to discuss both of them independently and more fully at some other occasion.

Chapter III gives some explicit constructions of factors, generalizing those of (**1**), pp. 192–204. And then, with the help of the above-mentioned tools of Chapters I, II, it is shown in Chapter IV, to which classes the factors of Chapter III belong. At this stage examples of factors of class (III_∞) are also obtained.

3. The main questions concerning factors were stated in (**1**) as Problems 1–11. Of these, Problems 1, 2, 5, and parts of 3, 4, 6, 7, 8, 9, 10 were answered in (**1**); Problem 11 was answered in (**2**). Our present result will settle the remainder of Problems 3, 4. Important parts of 6, 7, 8, 9, 10 remain unsettled. Especially

the following question is still open, and it is in our opinion of a quite decisive character: Are all factors of class (II_1) isomorphic to each other? (Part of 6.) We think that the answer to this question will greatly affect the applicability of the theory to factors, especially to quantum mechanics.

F. J. Murray and the author believe that the answer is negative, and that further invariants—probably of an algebraical and group theoretical nature—exist.

Certain partial results have been obtained, and they will be discussed elsewhere. But the main question (cf. above) is still unanswered.

4. The work **(7)** of the author also led to factors. The parts of the ring $\mathfrak{C}^{*1}$ in the various incomplete direct products $\prod^{\mathfrak{D}}_{\otimes n=1,2,\cdots}(\mathfrak{H}_{(n,1)} \otimes \mathfrak{H}_{(n,2)})$ ((7), p. 71, Lemma 7.3.1) were found to be in certain cases factors of classes (I_∞), (II_1), and (II_∞). (Cf. ibid., pp. 71–77, in particular Lemmas 7.4.1, 7.5.1.) It is not difficult to show that the above-mentioned parts are always (that is, in every $\prod^{\mathfrak{D}}_{\otimes n=1,2,\cdots}(\mathfrak{H}_{(n,1)} \otimes \mathfrak{H}_{(n,2)})$ factors.

Application of our present results, namely of our Theorem IX, shows that these factors are of class (III_∞) in certain cases.

More precisely: It was shown in **(7)**, p. 71, Lemma 7.3.1, that the part of $\mathfrak{C}^{*1}$ in $\prod^{\mathfrak{D}}_{\otimes n=1,2,\cdots}(\mathfrak{H}_{(n,1)} \otimes \mathfrak{H}_{(n,2)})$ depended essentially (that is, up to isomorphisms) only upon a certain sequence $\alpha_1, \alpha_2, \cdots$ of constants $\geqq 0, \leqq 1$. It was also shown ibid., pp. 71–77, that its class was (I_∞) for $\alpha_1 = \alpha_2 = \cdots = 1$, (II_1) for $\alpha_1 = \alpha_2 = \cdots = 0$, and stated that it was (II_∞) for $\alpha_1 = \alpha_3 = \cdots = 0$, $\alpha_2 = \alpha_4 = \cdots = 1$. Now we can show this: The class is (III_∞), if for some $\delta > 0$ we have for infinitely many $n = 1, 2, \cdots$, $\delta \leqq \alpha_n \leqq 1 - \delta$.

The factors obtained by the above process have various instructive aspects, and will be discussed (together with the above-stated results) elsewhere.

The surmises stated at the end of **(7)** have to be modified in accordance with the results stated above.

5. For an account of modern operator theory in general, the reader is referred to **(3)** or **(4)**. For the other topics touched, a general orientation may be obtained from **(5)**, **(1)**. A familiarity with the methods and results of these papers will be assumed.

Our notations are the same as in the papers enumerated above, for a detailed description cf. **(1)**, pp. 126–127.

A detailed table of contents and a bibliography follow.

CONTENTS

INTRODUCTION

1. Statement of the main result 94
2. Norm, normedness, $A^{|E_1|E_2|\cdots}$ 94
3. The Problems 1–11 of **(1)** 94

4. Relation with (7) 95
5. Literature and notations 95

CONTENTS 95

BIBLIOGRAPHY 96

CHAPTER I. THE RELATIVE TRACE IN AN ARBITRARY FACTOR 97
1.1. Generalities 97
1.2. Definition and properties of $\overline{T}_{\mathfrak{M}}(A)$. ($A$ of finite rank)—Theorem I 97
1.3. Definition and properties of $[[A]]$. (A of finite rank)—Theorem II. 101
1.4. Fundamental sequences, normed A's, $[[A]]$ for normed A's—Theorem III 105
1.5. Properties of normed A's, of $[[A]]$ for normed A's. Projections—Theorems IV–V 112
1.6. Meaning of normedness and norm in each class (I_n)–(III_∞) 115

CHAPTER II. MAXIMUM ABELIAN RINGS AND THE OPERATION $A^{|E_1|E_2|\cdots}$ 118
2.1. Construction and properties of $A^{|E_1|E_2|\cdots}$.—Theorem VI 118
2.2. Criterion of $\mathfrak{M}$ being of class (III_∞) 122

CHAPTER III. CONSTRUCTION OF THE RINGS 124
3.1. Generalities 124
3.2. S, Γ, μ, $\mathfrak{H}_S = \mathfrak{H}_{S,\Gamma,\mu}$. The "differentiation-theorem"—Theorem VII 124
3.3. Relation of a group $\mathfrak{G}$ to S, Γ, μ. Being free, ergodic, measurable .. 131
3.4. $\mathfrak{H}^{\mathfrak{G}}$, $\mathfrak{H}_S^{\mathfrak{G}} = \mathfrak{H}_{S,\Gamma,\mu}^{\mathfrak{G}}$ 134
3.5. Constructions in $\mathfrak{H}_S$ 135
3.6. Constructions in $\mathfrak{H}_S^{\mathfrak{G}}$: $\mathfrak{M}$, $\mathfrak{M}'$ 139
3.7. First discussion of $\mathfrak{M}$, $\mathfrak{M}'$—Theorem VIII 145

CHAPTER IV. DETERMINATION OF THE CLASSES OF $\mathfrak{M}$, $\mathfrak{M}'$ 149
4.1. Generalities 149
4.2. $\mathfrak{G}$ measurable: Cases (I), (II) 150
4.3. $\mathfrak{G}$ non-measurable: Case (III)—Theorem IX 153
4.4. Examples of Case (III). Not coupled factors 158

BIBLIOGRAPHY

(1) F. J. MURRAY AND J. v. NEUMANN: "On rings of operators," *Annals of Mathematics*, vol. 37 (1936), pp. 116–229.

(2) F. J. MURRAY AND J. v. NEUMANN: "On rings of operators, II," *Transactions Amer. Math. Soc.*, vol. 41 (1937), pp. 208–248.

(3) M. H. STONE: *Linear transformations in Hilbert space*, Amer. Math. Soc. Colloquium Publications, vol. XV (1932).

(4) J. v. NEUMANN: "Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren," *Math. Annalen*, vol. 102 (1929), pp. 49–131.

(5) J. v. NEUMANN: "Zur Algebra der Funktionaloperatoren," *Math. Annalen*, vol. 102 (1929), pp. 370–427.
(6) J. v. NEUMANN: "Über Funktionen von Funktionaloperatoren," *Annals of Math.*, vol. 32 (1931), pp. 191–226.
(7) J. v. NEUMANN: "On infinite direct products," *Compositio Mathematica*, vol. 6 (1938), pp. 1–77.
(0) F. HAUSDORFF: *Mengenlehre*. (III[rd] Edition.) W. de Gruyter, Berlin (1935).

CHAPTER I. THE RELATIVE TRACE IN AN ARBITRARY FACTOR

1.1. Let $\mathfrak{H}$ be a Hilbert space, and $\mathfrak{M}$ a factor in $\mathfrak{H}$. The notation of the *relative trace* $T_{\mathfrak{M}}(A)$ was defined for all Hermitean $A \in \mathfrak{M}$ in (**1**), pp. 212–213, Definition 15.1.1, and for all $A \in \mathfrak{M}$ without restriction in (**2**), p. 218, Definition 2.2.1, provided that $\mathfrak{M}$ belongs to one of the *finite* classes: (I_n) $(n = 1, 2, \cdots)$, (II_1). We shall now extend this definition to all factors $\mathfrak{M}$ but (when $\mathfrak{M}$ is *infinite*) not all $A \in \mathfrak{M}$. We shall see, in particular, that when $\mathfrak{M}$ is *purely infinite*, i.e. of class (III_∞), our definition covers $A = 0$ only. Hence the factors $\mathfrak{M}$ for which we really achieve something, are those in the *not purely infinite* classes: (I_∞), (II_∞). (For (I_∞) in particular, we get the usual notion of the trace and the E. Schmidt class of operators in $\mathfrak{M}$ cf. the discussion (B) in §1.6 below. Nevertheless, we shall use this extended notion of the relative trace in this paper for a factor $\mathfrak{M}$ of unknown class, with the purpose of showing that $\mathfrak{M}$ is just of class (III_∞).)

1.2. Let $\mathfrak{M}$ be an arbitrary factor, $D(\mathfrak{M})$ its *relative dimension function*, in any normalization. ((**1**), p. 165, Definition 8.2.1, and p. 168, Theorem VII.) For any $A \in \mathfrak{M}$ the closed, linear set [Range A] can be formed, and is $\eta\, \mathfrak{M}$. (This notion is defined in (**1**), p. 141, Definition 4.2.1. If a proof of the statement is wanted, cf. (**1**), pp. 151–152, Lemmas 6.1.1, 6.2.1.)

DEFINITION 1.2.1. *For every $A \in M$ we define a "rank" as follows*:

$$\text{Rank } (A) = D([\text{Range } A]).$$

(The reader will easily verify that if $\mathfrak{M} = \mathfrak{B} =$ the ring of all bounded operators in $\mathfrak{H}$—which is of class (I_∞)—and if $D(\mathfrak{M})$ is chosen in the standard normalization—(**1**), p. 173, Lemma 8.6.1—then this is the matrix-rank in its usual sense.)

Together with $D(\mathfrak{M})$, Rank (A) can assume finite or infinite numerical values. When the former is the case, we say that A has *finite rank*.

LEMMA 1.2.1. Rank (A) *has the following properties*:

(i) Rank $(A) \geqq 0$, *it is* $= 0$ *if and only if* $A = 0$.
(ii) Rank (A) = Rank (A^*).
(iii) Rank (aA) = Rank (A) *for every complex number* $a \neq 0$.
(iv) Rank $(A + B) \leqq$ Rank (A) + Rank (B).
(v) Rank $(AB) \leqq$ Rank (A), Rank (B).
(vi) Rank $(E) = D(E)$ for every projection $E \in \mathfrak{M}$.

PROOF:

Ad (i): Rank $(A) = D([\text{Range } A])$ is clearly $\geqq 0$, and $= 0$ if and only if [Range A] = (0), i.e. Range A = (0), which means $A = 0$.

Ad (ii): Coincides with (**1**), p. 152, Lemma 6.2.1.

Ad (iii): Since A is linear, already Range (aA) = Range (A) if $a \neq 0$.

Ad (iv): Clearly Range $(A + B) \subset$ [Range A, Range B], hence [Range $(A + B)$] $\subset$ [[Range A], [Range B]], whence (**1**), p. 210, Lemma 14.2.3, gives the desired result.

Ad (v): Clearly Range $(AB) \subset$ Range A, hence [Range (AB)] $\subset$ [Range A], and so Rank $(AB) \leqq$ Rank (A). Combining this with (ii), we obtain: Rank (AB) = Rank $((AB)^*)$ = Rank $(B^*A^*) \leqq$ Rank (B^*) = Rank (B).

This completes the proof.

Ad (vi): If E is the projection of the closed linear set $\mathfrak{M}$, then Range $E = \mathfrak{M}$, hence [Range E] = $\mathfrak{M}$.

LEMMA 1.2.2. *The finiteness of rank has the following properties*:

(i) 0 *has finite rank.*

(ii) A *has a finite rank if and only if* A^* *has a finite rank.*

(iii) A *has a finite rank if and only if* aA *has a finite rank,* a *being any fixed complex number* $\neq 0$.

(iv) $A + B$ *has a finite rank, if the same is true for both* A *and* B.

(v) AB *has a finite rank, if the same is true for either* A *or* B.

(vi) E *has a finite rank if and only if it is finite in the sense of* (**1**), p. 155, Definition 7.1.1, E *being any fixed projection* $\epsilon \mathfrak{M}$.

PROOF: These (i)–(vi) are immediate consequences of the corresponding (i)–(vi) in Lemma 1.2.1.—

Throughout what follows we will make use of the results of these Lemmas 1.2.1, 1.2.2, without referring to them explicitly.

DEFINITION 1.2.2. *A closed, linear set* $\mathfrak{N} \eta \mathfrak{M}$—*or its projection* $E \epsilon \mathfrak{M}$ ((**1**), p. 141, *Lemma* 4.2.1)—"*contains*" *an* $A \epsilon \mathfrak{M}$ *if*

$$[\text{Range } A], \quad [\text{Range } A^*] \subset \mathfrak{N}.$$

LEMMA 1.2.3. E *contains* A *if and only if*

$$EA = AE = A.$$

PROOF: [Range A] $\subset \mathfrak{N}$ is equivalent to Range $A \subset \mathfrak{N}$, that is, to $Af \epsilon \mathfrak{N}$ for every $f \epsilon \mathfrak{H}$, that is to $EAf = Af$. In other words: To $EA = A$.

Similarly [Range A^*] $\subset \mathfrak{N}$ means $EA^* = A^*$, that is (applying * to both sides) $AE = A$.

Hence both together amount to $EA = AE = A$.

LEMMA 1.2.4. A $(\epsilon \mathfrak{M})$ *is of finite rank, if and only if it is contained in a finite* $\mathfrak{N}$ $(\eta \mathfrak{M})$.

PROOF: *Sufficiency*: [Range A] $\subset \mathfrak{N}$, $\mathfrak{N}$ is finite, hence [Range A] is finite, that is, Rank $(A) = D([\text{Range } A])$ is a finite number.

Necessity: A has finite rank, hence A^* has also finite rank, so [Range A], [Range A^*] are finite, and with them $\mathfrak{M} = [[\text{Range } A], [\text{Range } A^*]]$ ((**1**), p. 160, Lemma 7.3.5). This $\mathfrak{M}$ meets all requirements.—

Given a finite $\mathfrak{M}$ ($\eta\, \mathfrak{M}$)—and its projection E ($\epsilon\, \mathfrak{M}$)—then those $A \epsilon \mathfrak{M}$ which are contained in $\mathfrak{M}$ are by Lemma 1.2.3 precisely the elements which the correspondence $X \rightleftarrows X_{(\mathfrak{M})}$ (discussed in (**1**), pp. 186–187, Definition 11.3.1, Lemmas 11.3.2–11.3.4) maps on the factor $\mathfrak{M}_{(\mathfrak{M})}$ in $\mathfrak{M}$ (instead of $\mathfrak{H}$!). So we see: If A runs over all elements of $\mathfrak{M}$ which are contained in $\mathfrak{M}$, then $A_{(\mathfrak{M})}$ runs over all elements of $\mathfrak{M}_{(\mathfrak{M})}$.

The factor $\mathfrak{M}_{(\mathfrak{M})}$ necessarily belongs to one of the finite classes: (I_n) ($n = 1, 2, \cdots$), (II_1) ((**1**), p. 188, Lemma 11.3.7, or p. 189, Lemma 11.4.3). Hence there exists a unique relative trace in $\mathfrak{M}_{(\mathfrak{M})}$ defined for all $A_{\mathfrak{M}} \epsilon \mathfrak{M}_{(\mathfrak{M})}$: $T_{\mathfrak{M}_{(\mathfrak{M})}}(A_{(\mathfrak{M})})$. (For the definition of $T_{\mathfrak{M}_{(\mathfrak{M})}}(A_{\mathfrak{M}})$ cf. the beginning of §1.1, above.) We now define:

DEFINITION 1.2.3. *If $A \epsilon \mathfrak{M}$ and of finite rank, then define for every finite, closed, linear set $\mathfrak{M} \eta \mathfrak{M}$ which contains A (such $\mathfrak{M}$ exist by Lemma* 1.2.4)

$$\overline{T}_{\mathfrak{M}_{(\mathfrak{M})}}(A) = D(\mathfrak{M}) T_{\mathfrak{M}_{(\mathfrak{M})}}(A_{(\mathfrak{M})}).$$

(This $\overline{T}_{\mathfrak{M}_{(\mathfrak{M})}}(A)$ depends also on the normalization of $D(\mathfrak{M})$, but we do not need to indicate this explicitly.)

LEMMA 1.2.5. *The $\overline{T}_{\mathfrak{M}_{(\mathfrak{M})}}(A)$ of Definition* 1.2.3 *depends on A only, and not on $\mathfrak{M}$ (as long as A, $\mathfrak{M}$ are related as described there).*

PROOF: In other words: If $\mathfrak{M}$, $\mathfrak{N}$ are both finite, closed linear sets, both $\eta\, \mathfrak{M}$ and both containing A, then we claim that

$$\overline{T}_{\mathfrak{M}_{(\mathfrak{M})}}(A) = \overline{T}_{\mathfrak{M}_{(\mathfrak{N})}}(A). \tag{1}$$

Consider first the case where $\mathfrak{M} \subset \mathfrak{N}$. Consider the operators $A_{\mathfrak{M}} \epsilon \mathfrak{M}_{(\mathfrak{M})}$. Form the operator $A_{\mathfrak{M}|\mathfrak{N}}$ in $\mathfrak{N}$, which is reduced by $\mathfrak{N}$ and which coincides with $A_{\mathfrak{M}}$ in $\mathfrak{M}$ and with 0 in $\mathfrak{N} - \mathfrak{M}$. We claim that $A_{\mathfrak{M}|\mathfrak{N}} \epsilon \mathfrak{M}_{(\mathfrak{N})}$. Indeed: $A_{\mathfrak{M}} \epsilon \mathfrak{M}_{(\mathfrak{M})}$ implies $A_{\mathfrak{M}} = A_{(\mathfrak{M})}$ for an $A \epsilon \mathfrak{M}$. As $E = P_{\mathfrak{M}} \epsilon \mathfrak{M}$, we have $(EAE)_{(\mathfrak{M})} = A_{(\mathfrak{M})}$, $EAE \epsilon \mathfrak{M}$, so we may replace A by EAE, which coincides with 0 in $\mathfrak{H} - \mathfrak{M}$. That is: We may assume that A itself coincides with 0 in $\mathfrak{H} - \mathfrak{M}$. Hence $A_{\mathfrak{M}|\mathfrak{N}} = A_{(\mathfrak{N})} \epsilon \mathfrak{M}_{(\mathfrak{N})}$.

Observe further:

(i)* $I_{(\mathfrak{M})} = E_{(\mathfrak{M})}$ *is the unit in* $\mathfrak{M}$, *hence for* $\mathfrak{M}_{(\mathfrak{M})}$. *And* $I_{(\mathfrak{N})|\mathfrak{N}} = E_{(\mathfrak{N})}$.

(ii)* $(a \cdot A_{\mathfrak{M}})_{|\mathfrak{N}} = a \cdot A_{\mathfrak{M}|\mathfrak{N}}$.

(iii)* $(A_{\mathfrak{M}} + B_{\mathfrak{M}})_{|\mathfrak{N}} = A_{\mathfrak{M}|\mathfrak{N}} + B_{\mathfrak{M}|\mathfrak{N}}$.

(iv)* *If* $A_{\mathfrak{M}}$ *is definite, then* $A_{\mathfrak{M}|\mathfrak{N}}$ *is also definite.*

(v)* $(A^*_{\mathfrak{M}})_{|\mathfrak{N}} = A^*_{\mathfrak{M}|\mathfrak{N}}$.

(vi)* $(A_{\mathfrak{M}} B_{\mathfrak{M}})_{|\mathfrak{N}} = A_{\mathfrak{M}|\mathfrak{N}} B_{\mathfrak{M}|\mathfrak{N}}$.

Now put for every $A_{\mathfrak{M}} \epsilon \mathfrak{M}_{(\mathfrak{M})}$

$$t(A_{\mathfrak{M}}) = \frac{D(\mathfrak{N})}{D(\mathfrak{M})}\, T_{\mathfrak{M}_{(\mathfrak{N})}}(A_{\mathfrak{M}|\mathfrak{N}}). \tag{2}$$

Comparing (2) and (i)*-(vi)* with (i)-(vi) in "Property III" in (2), pp. 218-219, we see immediately that $t(A_{\mathfrak{M}})$ fulfills (ii)-(vi), including (iv)', mentioned there (for $\mathfrak{M}_{(\mathfrak{M})}$, $\mathfrak{M}$ in place of $\mathfrak{M}$, $\mathfrak{H}$). (vi)' is also fulfilled, since it is a consequence of (vi). (Put $U^{-1}A$, U for A, B.) Let us finally check (i). Considering (2) and (i)*, it becomes

$$1 = \frac{D(\mathfrak{M})}{D(\mathfrak{N})} T\mathfrak{M}_{(\mathfrak{N})}(E_{(\mathfrak{N})}), \quad \text{that is}$$

$$T\mathfrak{M}_{(\mathfrak{N})}(E_{(\mathfrak{N})}) = \frac{D(\mathfrak{N})}{D(\mathfrak{M})}. \tag{3}$$

Now we know from "Property III," (*), in (2), pp. 218-219, together with (1), p. 219 (end of the first paragraph), that we have for every idempotent

$$F = P_{\mathfrak{P}}, \qquad \mathfrak{P} \subset \mathfrak{N}, \qquad T\mathfrak{M}_{(\mathfrak{N})}(F_{(\mathfrak{N})}) = D'(F_{(\mathfrak{N})}) = D'(\mathfrak{P}),$$

where $D'(\mathfrak{P})$ is the relative dimension function of $\mathfrak{M}_{(\mathfrak{N})}$ in its standard normalization—that is, with $D'(\mathfrak{N}) = 1$. By (1), p. 188, Lemma 11.3.7, $D(\mathfrak{P})$ is also a relative dimension function of $\mathfrak{M}_{(\mathfrak{N})}$, but in order to make its normalization the standard one, we must divide it by $D(\mathfrak{N})$. So $D'(\mathfrak{P}) = D(\mathfrak{P})/D(\mathfrak{N})$, and consequently $T\mathfrak{M}_{(\mathfrak{N})}(F_{(\mathfrak{N})}) = D(\mathfrak{P})/D(\mathfrak{N})$. Now putting $\mathfrak{P} = \mathfrak{M}$, $F = E$ gives (3). Thus (i) is also established.

We have thus shown that $t(A_{\mathfrak{M}})$ fulfills all conditions (i)-(vi)' in "Property III" in (2), pp. 218-219. Hence by "Property IV" in (2), p. 219, we have $T\mathfrak{M}_{(\mathfrak{M})}(A_{\mathfrak{M}}) = t(A_{\mathfrak{M}})$. Or, if we substitute this in (2):

$$D(\mathfrak{M})T\mathfrak{M}_{(\mathfrak{M})}(A_{\mathfrak{M}}) = D(\mathfrak{N})T\mathfrak{M}_{(\mathfrak{N})}(A_{\mathfrak{M}|\mathfrak{N}}), \tag{4}$$

Consider now our original A. Since it is contained in $\mathfrak{M}$, so $EA = AE = A$ by Lemma 1.2.3, therefore it coincides with 0 in $\mathfrak{H} - \mathfrak{M}$. Therefore we have for $A_{\mathfrak{M}} = A_{(\mathfrak{M})}$ also $A_{\mathfrak{M}|\mathfrak{N}} = A_{(\mathfrak{N})}$. By this substitution, however, (4) become

$$D(\mathfrak{M})T\mathfrak{M}_{(\mathfrak{M})}(A_{(\mathfrak{M})}) = D(\mathfrak{N})T\mathfrak{M}_{(\mathfrak{N})}(A_{(\mathfrak{N})}),$$

which, by Definition 1.2.3, is precisely (1).

This completes the proof for the case where $\mathfrak{M} \subset \mathfrak{N}$.

Consider next the case where $\mathfrak{M}$, $\mathfrak{N}$ are arbitrary. Then form $\mathfrak{P} = [\mathfrak{M}, \mathfrak{N}]$. Along with $\mathfrak{M}$, $\mathfrak{N}$ this $\mathfrak{P}$ will also be $\eta\,\mathfrak{M}$, and it will also be finite ((1), p. 160, Lemma 7.3.5). We have $\mathfrak{M}$, $\mathfrak{N} \subset \mathfrak{P}$, and $\mathfrak{P}$ contains A since $\mathfrak{M}$, $\mathfrak{N}$ do. So we have (1) for $\mathfrak{M}$, $\mathfrak{P}$ and for $\mathfrak{N}$, $\mathfrak{P}$ (in place of $\mathfrak{M}$, $\mathfrak{N}$), and hence for $\mathfrak{M}$, $\mathfrak{N}$ too.

Thus the proof is completed.

DEFINITION 1.2.4. *Considering Lemma* 1.2.5 *we can drop the* $\mathfrak{M}$ *from the* $\overline{T}\mathfrak{M}_{(\mathfrak{M})}(A)$ *of Definition* 1.2.3. *We shall therefore write for every* $A \in \mathfrak{M}$ *of finite rank (in the sense of Definition* 1.2.3):

$$\overline{T}\mathfrak{M}(A) = \overline{T}\mathfrak{M}_{(\mathfrak{M})}(A).$$

We obtain now with ease:

THEOREM I. *$\overline{T}_{\mathfrak{M}}(A)$ has the properties* (ii)-(vi) *in "Property III" in* (**2**), pp. 218–219). *That is*:

(ii) $\overline{T}_{\mathfrak{M}}(aA) = a\overline{T}_{\mathfrak{M}}(A)$, *a a complex number.*

(iii) $\overline{T}_{\mathfrak{M}}(A + B) = \overline{T}_{\mathfrak{M}}(A) + \overline{T}_{\mathfrak{M}}(B)$.

(iv) $\overline{T}_{\mathfrak{M}}(A) \geqq 0$ *if A is definite.*

(iv)′ $\overline{T}_{\mathfrak{M}}(A) > 0$ *if A is definite and* $\neq 0$.

(v) $\overline{T}_{\mathfrak{M}}(A^*) = \overline{\overline{T}_{\mathfrak{M}}(A)}$.

(vi) $\overline{T}_{\mathfrak{M}}(AB) = \overline{T}_{\mathfrak{M}}(BA)$ *if either A or B has a finite rank, and both are* $\epsilon \mathfrak{M}$.

(vi)′ $\overline{T}_{\mathfrak{M}}(U^{-1}AU) = \overline{T}_{\mathfrak{M}}(A)$ *if* U $(\epsilon \mathfrak{M})$ *is unitary.*

We have also:

(vii) $\overline{T}_{\mathfrak{M}}(E) = D(E)$ if E $(\epsilon \mathfrak{M})$ *is a finite projection.*

(Throughout (i)–(vii) remember Lemma 1.2.2.)

PROOF: Ad (ii), (iv), (iv)′, (v): They follow immediately from the corresponding statements in "Property III" in (**2**), pp. 218–219. We must only observe for (ii), (v), that $\mathfrak{M}$ contains aA, A^* along with A. And for (iv)′, that if $\mathfrak{M}$ contains A, then $A_{(\mathfrak{M})} = 0$ implies $A = 0$.

Ad (iii) and (vi), when A, B have both finite rank: In order to deduce them from the corresponding statements of "Property III," as above, we only need a finite $\mathfrak{M} \eta \mathfrak{M}$ which contains both A, B. Let $\mathfrak{M}$, $\mathfrak{N} \eta \mathfrak{M}$ be finite and contain A, B respectively. Then $\mathfrak{P} = [\mathfrak{M}, \mathfrak{N}]$ meets all requirements. (For the finiteness by (**1**), p. 160, Lemma 7.3.5.)

Ad (vi): Since (vi) is symmetric in A, B, we may assume that A has a finite rank. So has BA, choose a finite $\mathfrak{M} \eta \mathfrak{M}$ which contains both A, BA (cf. above). So we have for $E = P_{\mathfrak{M}}$ by Lemma 1.2.3, $AE = A$, $EBA = BA$. Consequently

$$AB = AE \cdot B = A \cdot EB,$$

$$BA = EBA \quad = EB \cdot A.$$

Now A has a finite rank, and EB has also a finite rank along with E. So our previous result gives $\overline{T}_{\mathfrak{M}}(A \cdot EB) = \overline{T}_{\mathfrak{M}}(EB \cdot A)$, and consequently $\overline{T}_{\mathfrak{M}}(AB) = \overline{T}_{\mathfrak{M}}(BA)$.

Ad (vi)′: Put $U^{-1}A$, U for A, B in (vi). $U^{-1}A$ has a finite rank along with A.

Ad (vii): Put $E = P_{\mathfrak{M}}$, $\mathfrak{M}$ is finite along with E, it is $\eta \mathfrak{M}$, and obviously contains E. $E_{(\mathfrak{M})} = I_{(\mathfrak{M})}$ is the unit in $\mathfrak{M}$, hence for $\mathfrak{M}_{(\mathfrak{M})}$. So we obtain from Definition 1.2.3, considering (i) in "Property III," as above:

$$\overline{T}_{\mathfrak{M}}(E) = D(\mathfrak{M}) T_{\mathfrak{M}_{(\mathfrak{M})}}(E_{(\mathfrak{M})}) = D(\mathfrak{M}).$$

1.3. Our new relative trace $\overline{T}_{\mathfrak{M}}(A)$ was only defined for the $A \epsilon \mathfrak{M}$ with a finite rank. We now wish to extend its definition to a wider domain of operators $A \epsilon \mathfrak{M}$, with the help of a suitable limiting process.

We first introduce a *norm* in the space of all $A \in \mathfrak{M}$ of finite rank, to be denoted by $[[A]]$, which is analogous to the notion of (2), p. 241, Definition 4.3.1, and Lemma 4.3.2 (especially the latter).

DEFINITION 1.3.1. *If $A \in \mathfrak{M}$ has a finite rank, then we define the "norm" of A, a number $[[A]] \geqq 0$, by*

$$[[A]] = (\overline{T}_{\mathfrak{M}}(A^*A))^{\frac{1}{2}} = (\overline{T}_{\mathfrak{M}}(AA^*))^{\frac{1}{2}}.$$

(Remember Theorem I, (iv), (vi). A^*A, AA^* are obviously definite.)

THEOREM II. *$[[A]]$ has the characteristic properties of a norm in a linear space. That is:*

(i) *Always $[[A]] \geqq 0$, and $= 0$ if and only if $A = 0$.*

(ii) $[[aA]] = |a| \cdot [[A]]$ *for every complex a.*

(iii) $[[A + B]] \leqq [[A]] + [[B]]$.

$[[A]]$ also possesses the following further properties:

(iv) $[[A^*]] = [[A]]$.

(v) $[[AB]] \begin{cases} \leqq |||A||| \cdot [[B]], & \text{where only } B \text{ need be of finite rank,} \\ \leqq |||B||| \cdot [[A]], & \text{where only } A \text{ need be of finite rank.} \end{cases}$

(vi) $|\overline{T}_{\mathfrak{M}}(AB)| \leqq [[A]] \cdot [[B]]$, *where A, B must both be of finite rank.*

(vii) $[[E]] = (D(E))^{\frac{1}{2}}$, *where E is a finite projection.*

PROOF: We prove these statements in a somewhat changed order:

Ad (i): Immediately by Theorem I, (iv), (iv)′ if we recall that $A^*A = 0$ implies $A = 0$.[1]

Ad (ii): Immediately by Theorem I, (ii).

Ad (iv): Immediately by the form of Definition 1.3.1.

Ad (vii): Immediately by Theorem I, (vii), since $E^*E = E^2 = E$.

Ad (v), first inequality: Put $|||A||| = a$, then $a^2I - A^*A$ is definite,[2] hence $B^*(a^2I - A^*A)B$ is definite.[3] Now

$$B^*(a^2I - A^*A)B = a^2B^*B - B^*A^*AB = a^2B^*B - (AB)^*(AB),$$

and so this expression is definite too. Therefore Theorem I, (iv), gives, together with (ii), (iii), eod.:

$$0 \leqq \overline{T}_{\mathfrak{M}}(a^2B^*B - (AB)^*(AB))$$

$$= a^2\overline{T}_{\mathfrak{M}}(B^*B) - \overline{T}_{\mathfrak{M}}((AB)^*(AB)) = a^2[[B]]^2 - [[AB]]^2,$$

$$[[AB]] \leqq a[[B]] = |||A||| \cdot [[B]].$$

Ad (v), second inequality: Combine the above result with (iv):

$$[[AB]] = [[(AB)^*]] = [[B^*A^*]] \leqq |||B^*||| \cdot [[A^*]] = {}^{4}|||B||| \cdot [[A]]$$

---

[1] $||Af||^2 = (Af, Af) = (A^*Af, f) = 0$, $||Af|| = 0$, hence $A = 0$.

[2] $((a^2I - A^*A)f, f) = a^2(f, f) - (A^*Af, f) = a^2(f, f) - (Af, Af) = a^2||f||^2 - ||Af||^2 \geqq 0$.

[3] B^*CB is always definite along with C:

$$(B^*CBf, f) = (CBf, Bf) \geqq 0.$$

[4] $|||B^*||| = |||B|||$, cf. (5), p. 384, footnote 36.

Ad (vi): Considering (iv), we might as well prove

$$|\ \overline{T}_{\mathfrak{M}}(A^*B)\ | \leqq [[A]]\cdot[[B]]. \tag{5}$$

In what follows we shall make use continuously of Theorem I, (ii), (iii), (v), and of (ii) above:

$$\begin{aligned} 0 \leqq [[A - B]]^2 &= \overline{T}_{\mathfrak{M}}((A - B)^*(A - B)) \\ &= \overline{T}_{\mathfrak{M}}(A^*A + B^*B - A^*B - B^*A) \\ &= \overline{T}_{\mathfrak{M}}(A^*A) + \overline{T}_{\mathfrak{M}}(B^*B) - \overline{T}_{\mathfrak{M}}(A^*B) - \overline{T}_{\mathfrak{M}}(B^*A) \\ &= [[A]]^2 + [[B]]^2 - \overline{T}_{\mathfrak{M}}(A^*B) - \overline{\overline{T}_{\mathfrak{M}}(A^*B)} \\ &= [[A]]^2 + [[B]]^2 - 2\Re\overline{T}_{\mathfrak{M}}(A^*B), \end{aligned}$$

$$\Re\overline{T}_{\mathfrak{M}}(A^*B) \leqq \tfrac{1}{2}[[A]]^2 + \tfrac{1}{2}[[B]]^2.$$

Replace A, B by aA, $\frac{1}{a}B$, where $a > 0$; then this becomes

$$\Re\overline{T}_{\mathfrak{M}}(A^*B) \leqq \frac{a^2}{2}[[A]] + \frac{1}{2a^2}[[B]]^2.$$

The greatest lower bound of the right side for all $a > 0$ is $[[A]]\cdot[[B]]$, hence

$$\Re\overline{T}_{\mathfrak{M}}(A^*B) \leqq [[A]]\cdot[[B]].$$

Now replace B by $e^{i\alpha}B$, α real, then this becomes

$$\Re(e^{i\alpha}\overline{T}_{\mathfrak{M}}(A^*B)) \leqq [[A]]\cdot[[B]].$$

The maximum of the left side for all real α is $|\ \overline{T}_{\mathfrak{M}}(A^*B)\ |$, hence (5) results.
Ad (iii): Use Theorem I, (iii), and (iv), (vi) above:

$$\begin{aligned} [[A + B]]^2 &= \overline{T}_{\mathfrak{M}}((A + B)^*(A + B)) = \overline{T}_{\mathfrak{M}}(A^*A + B^*B + A^*B + B^*A) \\ &= \overline{T}_{\mathfrak{M}}(A^*A) + \overline{T}_{\mathfrak{M}}(B^*B) + \overline{T}_{\mathfrak{M}}(A^*B) + T_{\mathfrak{M}}(B^*A) \\ &\leqq [[A]]^2 + [[B]]^2 + 2[[A]]\cdot[[B]] = ([[A]] + [[B]])^2, \end{aligned}$$

$$[[A + B]] \leqq [[A]] + [[B]].$$

Throughout what follows we shall make use of the results of these Theorems I and II without referring to them explicitly. Theorem II, (i)–(iii), give us also the right to use $[[A]]$ as a *norm* and $[[A - B]]$ as a *distance*, without any further comment.

In $|||A|||$ and $[[A]]$ we have two different numerical evaluations for the *size* of A. It is therefore desirable to evaluate each one in terms of the other. This is only feasible to a very limited extent, and is expressed by the Lemmas 1.3.1, 1.3.2, which follow. (Cf., however, B.3 in (B) in §1.6, when $\mathfrak{M}$ is in class (I_∞).)

LEMMA 1.3.1. $[[A]] \leqq (\text{Rank } (A))^{\frac{1}{2}} \cdot |||A|||$.

PROOF: Put $\mathfrak{M} = [\text{Range } A]$, $E = P_{\mathfrak{M}}$. Then $A \epsilon \mathfrak{M}$ implies $\mathfrak{M} \eta \mathfrak{M}$, $E \epsilon \mathfrak{M}$ (cf. Definition 1.2.2), and $D(E) = D(\mathfrak{M}) = D([\text{Range } A]) = \text{Rank } (A)$, $[[E]] = (D(E))^{\frac{1}{2}} = (\text{Rank } (A))^{\frac{1}{2}}$. Every Af belongs to $\mathfrak{M}$ so $EAf = Af$, that is $EA = A$. Hence

$$[[A]] = [[EA]] \leqq [[E]] \cdot |||A||| = (\text{Rank } (A))^{\frac{1}{2}} \cdot |||A|||.$$

LEMMA 1.3.2. *If* $[[A]] \leqq \delta\epsilon$ *with* $\delta, \epsilon > 0$, *then there exists an* $E \epsilon \mathfrak{M}$ *with* $D(E) \leqq \epsilon^2$, *so that* $|||A(I - E)||| \leqq \delta$.

PROOF: A^*A is Hermitean and bounded, so it has a spectral form

$$(A^*Af, g) = \int_{-\infty}^{\infty} \lambda d\,(E(\lambda)f, g. \tag{6}$$

(This is an application of the well-known spectral theorem for bounded operators, cf. e.g. (**5**), pp. 389–390, footnote 42, and p. 418, or (**1**), p. 212, Definition 15.1.1. The Stieltjes integral converges, cf. eod.) $E(\lambda)$, $-\infty < \lambda < +\infty$, is a *resolution of unity* (cf. eod.). Put $f = g$ in (6), since $(A^*Af, f) = (Af, Af) = ||Af||^2$, $(E(\lambda)f, f) = ||E(\lambda)f||^2$, we then obtain

$$||Af||^2 = \int_{-\infty}^{\infty} \lambda\, d(||E(\lambda)f||^2). \tag{7}$$

For $\lambda_0 < 0$ put $f = E(\lambda_0)g$. Then the right side in (7) becomes

$$= \int_{-\infty}^{\lambda_0} \lambda\, d(||E(\lambda)g||^2) \leqq \lambda_0 \int_{-\infty}^{\lambda_0} d(||E(\lambda)g||^2) = \lambda_0 ||E(\lambda_0)g||^2.$$

The left side of (7) is $\geqq 0$, hence $\lambda_0 < 0$ gives $||E(\lambda_0)g||^2 \leqq 0$, so $||E(\lambda_0)g|| = 0$, $E(\lambda_0)g = 0$. That is:

$$\lambda < 0 \quad \text{implies} \quad E(\lambda) = 0. \tag{8}$$

Now (7), (8) give for $f = E(\delta^2)g$:

$$||AE(\delta^2)g||^2 = \int_0^{\delta^2} \lambda\, d(||E(\lambda)g||^2) \leqq \delta^2 \int_0^{\delta^2} d(||E(\lambda)g||^2)$$

$$= \delta^2 ||E(\delta^2)g||^2 \leqq \delta^2 ||g||^2,$$

$$||AE(\delta^2)g|| \leqq \delta ||g||,$$

that is

$$|||AE(\delta^2)||| \leqq \delta. \tag{9}$$

On the other hand we have by (7), (8)

$$(\{A^*A - \delta^2(I - E(\delta^2))\}f, f) = (A^*Af, f) - \delta^2 \| (I - E(\delta^2))f \|^2$$

$$= \int_0^\infty \lambda \, d(\| E(\lambda)f \|^2) - \delta^2 \int_{\delta^2}^\infty d(\| E(\lambda)f \|^2)$$

$$\geqq \int_{\delta^2}^\infty \lambda \, d(\| E(\lambda)f \|^2) - \delta^2 \int_{\delta^2}^\infty d(\| E(\lambda)f \|^2)$$

$$= \int_{\delta^2}^\infty (\lambda - \delta^2) \, d(\| E(\lambda)f \|^2) \geqq 0.$$

So $A^*A - \delta^2(I - E(\delta^2))$ is definite, hence

$$[[A]]^2 = \overline{T}_{\mathfrak{M}}(A^*A) \geqq \delta^2 \overline{T}_{\mathfrak{M}}(I - E(\delta^2)) = \delta^2 D(I - E(\delta^2)).$$

But $[[A]]^2 \leqq \delta^2\epsilon^2$, consequently

$$D(I - E(\delta^2)) \leqq \epsilon^2. \tag{10}$$

So (9), (10) state that $E = I - E(\delta^2)$ meets all our requirements.

1.4. DEFINITION 1.4.1. *A sequence of operators $A_1, A_2, \cdots$ is called "regular" if it possesses the following properties:*

(i) *All $A_n \in \mathfrak{M}$.*

(ii) *All A_n have finite ranks. (But their ranks need not be bounded.)*

(iii) *The numbers $|||A_1|||, |||A_2|||, \cdots$ are bounded.*

DEFINITION 1.4.2. *A regular sequence $A_1, A_2, \cdots$ is "fundamental" if*

$$\lim_{m,n\to\infty} [[A_m - A_n]] = 0.$$

DEFINITION 1.4.3. *Two regular sequences $A_1, A_2, \cdots$ and $B_1, B_2, \cdots$ are "equivalent" if*

$$\lim_{n\to\infty} [[A_n - B_n]] = 0.$$

In all our considerations about operator-convergence which will follow, we shall use *strong convergence*. This is the definition (cf. (**5**), pp. 381–382):

DEFINITION 1.4.4. *A sequence $A_1, A_2, \cdots$ is "strongly convergent" to the limit A, in symbols*

$$\lim_{n\to\infty} \text{str } A_n = A,$$

if these conditions are fulfilled:

(i) *The numbers $|||A_1|||, |||A_2|||, \cdots$ are bounded.*

(ii) *For every $f \in \mathfrak{H}$*

$$\lim_{m,n\to\infty} \| A_n f - Af \| = 0.$$

We shall use only the most obvious properties of strong convergence. (Cf. (5), pp. 382–384. In Definition 1.4.4 above, (i) is a consequence of (ii)—cf. (5), p. 382, footnote 35—but we do not need this fact.)

We proceed now to establish connections between fundamental sequences and equivalence on one hand, and strong convergence on the other.

LEMMA 1.4.1. *Let $A_1, A_2, \cdots$ be a regular sequence with*

$$\lim_{n\to\infty} [[A_n]] = 0.$$

Then it possesses a subsequence $A_{i_1}, A_{i_2}, \cdots$ $(i_n \to \infty)$ with

$$\lim_{n\to\infty} \operatorname{str} A_{i_n} = 0.$$

PROOF: For every $m = 1, 2, \cdots$ there exists an $n_0 = n_0(m)$, such that $n \geqq n_0$ implies $[[A_n]] \leqq 1/4^m$. Choose $i_1 < i_2 < \cdots$ with $i_m \geqq n_0(m)$. Then we have

$$[[A_{i_m}]] \leqq \frac{1}{4^m}.$$

Apply Lemma 1.3.2 with $\delta = \epsilon = 1/2^m$, then an $E_m \in \mathfrak{M}$ obtains, with

$$D(E_m) \leqq \frac{1}{4^m}, \tag{11}$$

$$\| A_{i_m}(I - E_m) \| \leqq \frac{1}{2^m}. \tag{12}$$

Put $E_m = P_{\mathfrak{M}_m}$, $\mathfrak{N}_m = [\mathfrak{M}_m, \mathfrak{M}_{m+1}, \cdots]$, $F_m = P_{\mathfrak{N}_m}$. Then we have also $F_m \in \mathfrak{M}$, and (1), p. 168, Lemma 8.3.2 gives

$$D(\mathfrak{N}_m) \leqq \sum_{n=m}^{\infty} D(\mathfrak{M}_n),$$

$$D(F_m) \leqq \sum_{n=m}^{\infty} D(E_n) \leqq \sum_{n=m}^{\infty} \frac{1}{4^n} = \frac{1}{3 \cdot 4^{m-1}},$$

$$D(F_m) \leqq \frac{1}{3 \cdot 4^{m-1}}. \tag{13}$$

Also, by construction of the $\mathfrak{N}_m$, $\mathfrak{N}_m \supset \mathfrak{N}_{m+1}$, $F_m \geqq F_{m+1}$, that is

$$\mathfrak{N}_1 \supset \mathfrak{N}_2 \supset \cdots, \qquad F_1 \geqq F_2 \geqq \cdots. \tag{14}$$

Hence by (4), p. 77, Theorem 19, we have $\lim \operatorname{str}_{m\to\infty} F_m = F = P_{\mathfrak{N}_1 \cdot \mathfrak{N}_2 \cdot \ldots}$ and by (1), p. 169, Lemma 8.3.4,

$$D(P_{\mathfrak{N}_1 \cdot \mathfrak{N}_2 \cdot \ldots}) = \lim_{m\to\infty} D(\mathfrak{N}_m), \qquad D(F) = \lim_{m\to\infty} D(F_m).$$

This is $= 0$ by (13), so $F = 0$, that is

$$\lim_{n\to\infty} \operatorname{str} F_m = 0. \tag{15}$$

(12) states $\| A_{i_m}(I - E_m)f \| \leqq 1/2^m \| f \|$, hence for $f \epsilon \mathfrak{H} - \mathfrak{M}_m$

$$\| A_{i_m} f \| \leqq \frac{1}{2^m} \| f \|. \tag{16}$$

By construction of $\mathfrak{N}_m$ $\mathfrak{M}_m \subset \mathfrak{N}_m$, hence $\mathfrak{H} - \mathfrak{M}_m \supset \mathfrak{H} - \mathfrak{N}_m$, and so (16) holds for all $f \epsilon \mathfrak{H} - \mathfrak{N}_m$. Assume finally $f \epsilon \mathfrak{H} - \mathfrak{N}_p$, $m \geqq p$. Then (14) gives $\mathfrak{N}_p \supset \mathfrak{N}_m$ so $\mathfrak{H} - \mathfrak{N}_p \subset \mathfrak{H} - \mathfrak{N}_m$, and so (16) holds again.

So if $f \epsilon \mathfrak{H} - \mathfrak{N}_p$, then (16) holds for all $m \geqq p$. That is: Then

$$\lim_{m \to \infty} \| A_{i_m} f \| = 0. \tag{17}$$

As (17) holds in every $\mathfrak{H} - \mathfrak{N}_p$ and considering (15), we see that the domain of validity of (17) is an everywhere dense set in $\mathfrak{H}$.[5]

But the domain of validity of (17) is a closed set, because the $A_1, A_2, \cdots$ are uniformly continuous by (iii) in Definition 1.4.1. Hence it comprises all $\mathfrak{H}$. Thus (ii) in Definition 1.4.4 is established for the sequence $A_{i_1}, A_{i_2}, \cdots$. (i) eod. follows from (iii) in Definition 1.4.1.

LEMMA 1.4.2. *Let $A_1, A_2, \cdots$ be a fundamental sequence. Then* $\lim \mathrm{str}_{n \to \infty} A_n$ *exists.*

PROOF: Since (i) in Definition 1.4.4 follows from (iii) in Definition 1.4.1, we must only prove the validity of (ii) in Definition 1.4.4 for a suitable A. That is: For every $f \epsilon \mathfrak{H}$ the existence of a $g \epsilon \mathfrak{H}$ (which then is to be Af) with $\lim_{n \to \infty} \| A_n f - g \| = 0$. And since $\mathfrak{H}$ is complete, this is certainly the case if

$$\lim_{m,\, n \to \infty} \| A_n f - A_n f \| = 0. \tag{18}$$

Therefore assume the opposite: That (18) is not true for a suitable $f \epsilon \mathfrak{H}$. So we have an $f_0 \epsilon \mathfrak{H}$ and a $\delta > 0$, such that

$$\| A_{i_n} f_0 - A_{j_n} f_0 \| \geqq \delta \quad \text{for} \quad n = 1, 2, \cdots, \quad \text{where} \quad i_n, j_n \to \infty. \tag{19}$$

From the fact that $A_1, A_2, \cdots$ is a fundamental sequence, we conclude immediately that $A_{i_1} - A_{j_2}, A_{i_2} - A_{j_2}, \cdots$ is a regular sequence with $\lim_{n \to \infty} [[A_{i_n} - A_{j_n}]] = 0$. Therefore we may apply Lemma 1.4.1 to $A_{i_1} - A_{j_1}$, $A_{i_2} - A_{j_2}, \cdots$ (in place of $A_1, A_2, \cdots$). This sequence possesses a subsequence $A_{i_{k_1}} - A_{j_{k_1}}, A_{i_{k_2}} - A_{j_{k_2}}, \cdots$ ($k_n \to \infty$), with $\lim \mathrm{str}_{n \to \infty} (A_{i_{k_n}} - A_{j_{k_n}}) = 0$. Thus in particular $\lim_{n \to \infty} \| (A_{i_{k_n}} - A_{j_{k_n}}) f_0 \| = 0$, that is

$$\lim_{n \to \infty} \| A_{i_{k_n}} f_0 - A_{j_{k_n}} f_0 \| = 0, \tag{20}$$

(19) and (20) contradict each other, and this completes the proof.

LEMMA 1.4.3. *Let $A_1, A_2, \cdots$ be a fundamental sequence with* $\lim \mathrm{str}_{n \to \infty} A_n = 0$. *Then it possesses a subsequence $A_{i_1}, A_{i_2}, \cdots$ ($i_n \to \infty$) with*

$$\lim_{n \to \infty} [[A_{i_n}]] = 0.$$

[5] Consider an aribtrary $f \epsilon \mathfrak{H}$. Then $f_p = (I - F_p) f \epsilon \mathfrak{H} - \mathfrak{N}_p$, and $\lim_{p \to \infty} \| f - f_p \| = \lim_{p \to \infty} \| F_p f \| = 0$ by (15).

PROOF: Since $A_1, A_2, \cdots$ is a fundamental sequence, we have $\lim_{m,n\to\infty} [[A_m - A_n]] = 0$. Hence there exists for every $p = 1, 2, \cdots$ an $n_0 = n_0(p)$ such that $m, n \geqq n_0$ imply $[[A_m - A_n]] \leqq 1/4^p$. Choose $i_1 < i_2 < \cdots$ with $i_p \geqq n_0(p)$. Then we have

$$[[A_{i_p} - A_{i_{p+1}}]] \leqq \frac{1}{4^p}.$$

Apply Lemma 1.3.2 with $\delta = \epsilon = 1/2^p$, then an $E_p \epsilon \mathfrak{M}$ obtains, with

$$D(E_p) \leqq \frac{1}{4^p}, \tag{21}$$

$$|||(A_{i_p} - A_{i_{p+1}})(I - E_p)||| \leqq \frac{1}{2^p}. \tag{22}$$

Let us now proceed as in the proof of Lemma 1.4.1. Put $E_p = P_{\mathfrak{M}_p}$, $\mathfrak{N}_p = [\mathfrak{M}_p, \mathfrak{M}_{p+1}, \cdots]$, $F_p = P_{\mathfrak{N}_p}$. Then we have again $F_m \epsilon \mathfrak{M}$, and as there

$$D(\mathfrak{N}_p) \leqq \sum_{q=p}^{\infty} D(\mathfrak{M}_q),$$

$$D(F_p) \leqq \sum_{q=p}^{\infty} D(E_q) \leqq \sum_{q=p}^{\infty} \frac{1}{4^q} = \frac{1}{3 \cdot 4^{q-1}},$$

$$D(F_p) \leqq \frac{1}{3 \cdot 4^{q-1}}. \tag{23}$$

By the construction of the $\mathfrak{N}_p$, $\mathfrak{N}_p \supset \mathfrak{N}_{p+1}$, so $\mathfrak{N}_1 \supset \mathfrak{N}_2 \supset \cdots$, and

$$\mathfrak{H} - \mathfrak{N}_1 \subset \mathfrak{H} - \mathfrak{N}_2 \subset \cdots. \tag{24}$$

Consider now an $f \epsilon \mathfrak{H} - \mathfrak{N}_q$. Then we have for every $p \geqq q$ also $f \epsilon \mathfrak{H} - \mathfrak{N}_p$ by (24), hence $(I - E_p)f = f$. So (22) gives for such an f

$$\| A_{i_p} f - A_{i_{p+1}} f \| \leqq \frac{1}{2^p} \| f \|.$$

If $r > s \geqq q$, then the above inequality may be applied to $p = s, s + 1, \cdots, r - 1$.

Adding gives

$$\| A_{i_s} f - A_{i_r} f \| = \left\| \sum_{p=s}^{r-1} A_{i_p} f - A_{i_{p+1}} f \right\|$$

$$\leqq \sum_{p=s}^{r-1} \| A_{i_p} f - A_{i_{p+1}} f \| \leqq \sum_{p=s}^{r-1} \frac{1}{2^p} \| f \| \leqq \frac{1}{2^{s-1}} \| f \|,$$

$$\| A_{i_s} f - A_{i_{r+1}} f \| \leqq \frac{1}{2^{s-1}} \| f \|. \tag{25}$$

Now $\lim \operatorname{str}_{n\to\infty} A_n = 0$ implies $\lim_{n\to\infty} \| A_n f \| = 0$, and a fortiori $\lim_{r\to\infty,\, r>s} \| A_{i_r} f \| = 0$. So if $s \geqq q$ is fixed, and we let $r \to \infty$ with $r > s$ then (25) gives:

$$\| A_{i_s} f \| \leqq \frac{1}{2^{s-1}} \| f \| \qquad \text{if } s \geqq q. \tag{26}$$

This is true for every $f \in \mathfrak{H} - \mathfrak{N}_q$ that is for every $f = (I - F_q)g$. Substituting this into (26) gives:

$$\| A_{i_s}(I - F_q)g \| \leqq \frac{1}{2^{s-1}} \| (I - F_q)g \| \leqq \frac{1}{2^{s-1}} \| g \|.$$

That is:

$$\| \| A_{i_s}(I - F_q) \| \| \leqq \frac{1}{2^{s-1}}. \tag{27}$$

Having obtained (27), we can now proceed to a general evaluation of the A_n. Since the numbers $\| \| A_1 \| \|$, $\| \| A_2 \| \|$, $\cdots$ are bounded, by (iii) in Definition 1.4.1, there exists a fixed number a with

$$\| \| A_n \| \| \leqq a \qquad \text{for } n = 1, 2, \cdots. \tag{28}$$

Hence (23), (28) give:

$$[[A_{i_s} F_q]] \leqq \| \| A_{i_s} \| \| \cdot [[F_q]] = \| \| A_{i_s} \| \| \cdot (D(F_q))^{\frac{1}{2}} \leqq \frac{a}{(3.4^{q-1})^{\frac{1}{2}}},$$

$$[[A_{i_s} F_q]] \leqq \frac{a}{(3.4^{q-1})^{\frac{1}{2}}}. \tag{29}$$

Consider now any finite projection $G \in \mathfrak{M}$. Then (27), (29) yield:

$$\begin{aligned}
[[GA_{i_s}]] &\leqq [[GA_{i_s} F_q]] + [[GA_{i_s}(I - F_q)]] \\
&\leqq \| \| G \| \| \cdot [[A_{i_s} F_q]] + \| \| A_{i_s}(I - F_q) \| \| \cdot [[G]] \\
&\leqq 1 \cdot \frac{a}{(3.4^{q-1})^{\frac{1}{2}}} + (D(G))^{\frac{1}{2}} \cdot \frac{1}{2^{s-1}} \\
&= \frac{a}{(3.4^{q-1})^{\frac{1}{2}}} + \frac{(D(G))^{\frac{1}{2}}}{2^{s-1}},
\end{aligned}$$

$$[[GA_{i_s}]] \leqq \frac{a}{(3.4^{q-1})^{\frac{1}{2}}} + \frac{(D(G))^{\frac{1}{2}}}{2^{s-1}}, \tag{30}$$

Put $\mathfrak{P}_q = [\text{Range } A_{i_q}]$, $G_q = P_{\mathfrak{P}_q}$. Again $\mathfrak{P}_q \eta \mathfrak{M}$, $G_q \in \mathfrak{M}$. And always $A_{i_q} f \in \mathfrak{P}_q$, $G_q A_{i_q} f = A_{i_q} f$, that is $G_q A_{i_q} = A_{i_q}$. So $(I - G_q)A_{i_q} = 0$, and consequently

$$(I - G_q)(A_{i_s} - A_{i_q}) = (I - G_q)A_{i_s}.$$

Hence, by using our original construction of the $i_1, i_2, \cdots$,

$$[[(I - G_q)A_{i_s}]] = [[(I - G_q)(A_{i_s} - A_{i_q})]]$$

$$\leqq |||\, I - G_q \,|||\cdot[[A_{i_s} - A_{i_q}]] \leqq 1\cdot\frac{1}{4^q} = \frac{1}{4^q},$$

$$[[(I - G_q)A_{i_s}]] \leqq \frac{1}{4^q} \quad \text{if } s \geqq q. \tag{31}$$

Now combine (30) with $G = G_q$ and (31), remembering that $D(G_q) = D(\mathfrak{P}_q) = \text{Rank } (A_{i_q})$. So we obtain:

$$[[A_{i_s}]] \leqq \frac{1}{4^q} + \frac{a}{(3\cdot 4^{q-1})^{\frac{1}{2}}} + \frac{(\text{Rank } (A_{i_q}))^{\frac{1}{2}}}{2^{s-1}} \qquad \text{if } s \geqq q. \tag{32}$$

Let finally an $\epsilon > 0$ be given. Choose a $q = q(\epsilon)$ so that $\frac{1}{4^q} + \frac{a}{(3\cdot 4^{q-1})^{\frac{1}{2}}} \leqq \frac{\epsilon}{2}$, and then an $s_0 = s_0(\epsilon, q)$ so, that $s_0 \geqq q$ and $\frac{(\text{Rank } (A_{i_q}))^{\frac{1}{2}}}{2^{s_0-1}} \leqq \frac{\epsilon}{2}$. Then (32) gives immediately for all $s \geqq s_0$ that $[[A_{i_s}]] \leqq \epsilon$. In other words:

$$\lim_{s\to\infty} [[A_{i_s}]] = 0.$$

Thus the proof is completed.—

Before we go on, we observe this: Both Lemmata 1.4.1, 1.4.3 could have been easily strengthened by replacing the subsequence $A_{i_1}, A_{i_2}, \cdots$ by the sequence $A_1, A_2, \cdots$ itself. But they suffice as they stand for our immediate purposes, and our ultimate results (Theorem III) yield those strengthened forms immediately.

THEOREM III. (i) *For every fundamental sequence* $A_1, A_2, \cdots$ $\lim \text{str}_{n\to\infty} A_n$ *exists, and belongs to* $\mathfrak{M}$.

(ii) *For two fundamental sequences* $A_1, A_2, \cdots$ *and* $B_1, B_2, \cdots$ *the equation*

$$\lim_{n\to\infty} \text{str } A_n = \lim_{n\to\infty} \text{str } B_n$$

holds if and only if the two sequences are equivalent.

(iii) *For every fundamental sequence* $A_1, A_2, \cdots$ $\lim_{n\to\infty} [[A_n]]$ (*numerical!*) *exists, and it depends on* $\lim \text{str}_{n\to\infty} A_n$ *only* (*and not on the sequence* $A_1, A_2, \cdots$ *itself!*).

PROOF: We prove these statements in a somewhat changed order:

Ad (i): *Existence of* $\lim \text{str}_{n\to\infty} A_n$: Identical with Lemma 1.4.2.

Ad (i): $\lim \text{str}_{n\to\infty} A_n \in \mathfrak{M}$: Clear, since $A_1, A_2, \cdots \in \mathfrak{M}$, and $\mathfrak{M}$ being a ring, is closed in the strong topology.

Ad (ii): *Sufficiency*: Let $A_1, A_2, \cdots$ and $B_1, B_2, \cdots$ be two equivalent fundamental sequences. Then $A_1, B_1, A_2, B_2, \cdots$ is also a fundamental sequence; denote it by $C_1, C_2, \cdots$. By (i) above $\lim \text{str}_{n\to\infty} C_n$ exists. Hence $\lim \text{str}_{n\to\infty} C_{2n-1} = \lim \text{str}_{n\to\infty} C_{2n}$, that is $\lim \text{str}_{n\to\infty} A_n = \lim \text{str}_{n\to\infty} B_n$.

Ad (iii): *Existence of* $\lim_{n\to\infty}[[A_n]]$: Since $A_1, A_2, \cdots$ is a fundamental sequence, we have $\lim_{m,n\to\infty}[[A_m - A_n]] = 0$. Now $|\,[[A_m]] - [[A_n]]\,| \leqq [[A_m - A_n]]$,[6] hence $\lim_{m,n\to\infty}([[A_m]] - [[A_n]]) = 0$. Therefore the (numerical) limit $\lim_{n\to\infty}[[A_n]]$ exists.

Ad (ii): *Necessity*: We may replace $A_1, A_2, \cdots$ and $B_1, B_2, \cdots$ by $A_1 - B_1$, $A_2 - B_2, \cdots$ and $0, 0, \cdots$. That is: We may assume, that $B_1 = B_2 = \cdots = 0$. So $A_1, A_2, \cdots$ is fundamental, and we assume that $\lim \mathrm{str}_{n\to\infty} A_n = 0$. Then Lemma 1.4.3 excludes that $\lim_{n\to\infty}[[A_n]]$ should exist and at the same time be $\neq 0$. But we proved just now that part of (iii) which states that $\lim_{n\to\infty}[[A_n]]$ does exist. Hence $\lim_{n\to\infty}[[A_n]] = 0$. In other words: $A_1, A_2, \cdots$ and $0, 0, \cdots$ are equivalent.

Ad (iii): $\lim \mathrm{str}_{n\to\infty} A_n$ determines $\lim_{n\to\infty}[[A_n]]$: We have two fundamental sequences $A_1, A_2, \cdots$ and $B_1, B_2, \cdots$ with $\lim \mathrm{str}_{n\to\infty} A_n = \lim \mathrm{str}_{n\to\infty} B_n$, and we want to prove that $\lim_{n\to\infty}[[A_n]] = \lim_{n\to\infty}[[B_n]]$. We proved just now the necessity of the criterion of (ii), hence $A_1, A_2, \cdots$ and $B_1, B_2, \cdots$ are at any rate equivalent. That is: $\lim_{n\to\infty}[[A_n - B_n]] = 0$. But $|\,[[A_n]] - [[B_n]]\,| \leqq [[A_n - B_n]]$ (cf.[6] above), hence $\lim_{n\to\infty}([[A_n]] - [[B_n]]) = 0$. Consequently $\lim_{n\to\infty}[[A_n]] = \lim_{n\to\infty}[[B_n]]$.

We can now define:

DEFINITION 1.4.5. *The fact that* $A_1, A_2, \cdots$ *is a fundamental sequence and that* $\lim \mathrm{str}_{n\to\infty} A_n = A$, *will be denoted by*

$$(A_1, A_2, \cdots) \sim A.$$

DEFINITION 1.4.6. *We shall say that an operator* A *is "normed" if a sequence* $A_1, A_2, \cdots$ *with* $(A_1, A_2, \cdots) \sim A$ *exists. By* (iii) *in Theorem* III, $\lim_{n\to\infty}[[A_n]]$ *exists for every such sequence, and it depends on* A *only (not on the sequence* $A_1, A_2, \cdots$ *itself!). We shall define the "norm" of* A, *the number* $[[A]]$, *as this limit*: $[[A]] = \lim_{n\to\infty}[[A_n]]$.

REMARKS: (i) Every A of finite rank is normed, since we then have $(A, A, \cdots) \sim A$. This makes clear also that in this case—i.e., when the old $[[A]]$ is also defined—the new $[[A]]$ coincides with the old one.

(ii) By (i) in Theorem III every normed A is $\epsilon\, \mathfrak{M}$.

(iii) By (i), (ii) in Theorem III the relation

$$(A_1, A_2, \cdots) \sim A$$

establishes a one-to-one correspondence between all equivalence-classes of mutually equivalent fundamental sequences $A_1, A_2, \cdots$ on one hand, and all normed A on the other.—

Throughout what follows we shall make use of the results of Theorem III. Definitions 1.4.5, 1.4.6, and of the above Remarks, without referring to them explicitly.

---

[6] Always $|\,[[A]] - [[B]]\,| \leqq [[A - B]]$. Proof: $[[A]] = [[B + (A - B)]] \leqq [[B]] + [[A - B]]$, $[[A]] - [[B]] \leqq [[A - B]]$. Interchanging A, B gives $[[B]] - [[A]] \leqq [[A - B]]$. Hence $\pm([[A]] - [[B]]) \leqq [[A - B]]$, that is $|\,[[A]] - [[B]]\,| \leqq [[A - B]]$.

1.5. LEMMA 1.5.1. (i) $(A_1, A_2, \cdots) \sim A$ *implies* $(aA_1, aA_2, \cdots) \sim aA$ *for every complex* a.

(ii) $(A_1, A_2, \cdots) \sim A$, $(B_1, B_2, \cdots) \sim B$ *imply* $(A_1 + B_1, A_2 + B_2, \cdots) \sim A + B$

(iii) $(A_1, A_2, \cdots) \sim A$ *implies* $(A_1B, A_2B, \cdots) \sim AB$ *and* $(BA_1, BA_2, \cdots) \sim BA$ *for every* $B \in \mathfrak{M}$.

(iv) $(A_1, A_2, \cdots) \sim A$ *is equivalent to* $(A_1^*, A_2^*, \cdots) \sim A^*$

PROOF: Ad (i): $aA_1, aA_2, \cdots$ is fundamental along with $A_1, A_2, \cdots$ because $[[aA_m - aA_n]] = [[a(A_m - A_n)]] = |a| \cdot [[A_m - A_n]]$. And $\lim \operatorname{str}_{n\to\infty} A_n = A$ implies $\lim \operatorname{str}_{n\to\infty} aA_n = aA$.

Ad (ii): $A_1 + B_1, A_2 + B_2, \cdots$ is fundamental along with $A_1, A_2, \cdots$ and $B_1, B_2, \cdots$ because $[[(A_m + B_m) - (A_n + B_n)]] = [[(A_m - A_n) + (B_m - B_n)]] \leqq [[A_m - A_n]] + [[B_m - B_n]]$. And $\lim \operatorname{str}_{n\to\infty} A_n = A$, $\lim \operatorname{str}_{n\to\infty} B_n = B$ imply $\lim \operatorname{str}_{n\to\infty} (A_n + B_n) = A + B$.

Ad (iii): $A_1B, A_2B, \cdots$ and $BA_1, BA_2, \cdots$ are fundamental along with $A_1, A_2, \cdots$ because

$$[[A_mB - A_nB]] = [[(A_m - A_n)B]] \leqq |||B||| \cdot [[A_m - A_n]],$$

$$[[BA_m - BA_n]] = [[B(A_m - A_n)]] \leqq |||B||| \cdot [[A_m - A_n]].$$

And $\lim \operatorname{str}_{n\to\infty} A_n = A$ implies $\lim \operatorname{str}_{n\to\infty} A_nB = AB$, $\lim \operatorname{str}_{n\to\infty} BA_n = BA$.

Ad (iv): Since ** is identity, it suffices to prove the forward implication. Now $A_1^*, A_2^*, \cdots$ is fundamental along with $A_1, A_2, \cdots$ because $[[A_m^* - A_n^*]] = [[(A_m - A_n)^*]] = [[A_m - A_n]]$. Hence a B with $\lim \operatorname{str}_{n\to\infty} A_n^* = B$ exists, and we know that we have $\lim \operatorname{str}_{n\to\infty} A_n = A$.[7] Hence we have for every $f \in \mathfrak{H}$ $\lim_{n\to\infty} ||A_nf - Af|| = 0$, $\lim_{n\to\infty} ||A_n^*f - Bf|| = 0$. So for all $f, g \in \mathfrak{H}$ $\lim_{n\to\infty} (A_nf, g) = (Af, g)$, $\lim_{n\to\infty} (A_n^*f, g) = (Bf, g)$.

Replacing f, g in the first relation by g, f and application of $\bar{}$ gives $\lim_{n\to\infty} (A_n^*f, g) = (A^*f, g)$. Hence $(A^*f, g) = (Bf, g)$ for all $f, g \in \mathfrak{H}$. Thus $A^* = B$, and consequently $\lim \operatorname{str}_{n\to\infty} A_n^* = A^*$. This completes the proof.—

It is now easy to prove the fundamental properties of the class of the normed A and of their norm $[[A]]$.

THEOREM IV. $[[A]]$ *(for normed* A*) has the characteristic properties of a norm in a linear space. That is*:

(i) *Always* $[[A]] \geqq 0$, *and* $= 0$ *if and only if* $A = 0$.

(ii) aA *is normed along with* A, *and* $[[aA]] = |a| \cdot [[A]]$ *for every complex* a.

(iii) $A + B$ *is normed along with* A, B *and* $[[A + B]] \leqq [[A]] + [[B]]$.

The following further statements are true:

(iv) A^* *is normed if and only if* A *is normed, and* $[[A^*]] = [[A]]$.

(v) AB *is normed if either* A *or* B *is normed (but both* $\in \mathfrak{M}$*), and*

$$[[AB]] \begin{cases} \leqq |||A||| \cdot [[B]] \text{ if } B \text{ is normed,} \\ \leqq |||B||| \cdot [[A]] \text{ if } A \text{ is normed.} \end{cases}$$

[7] $\lim \operatorname{str}_{n\to\infty}(A_n^*) = (\lim \operatorname{str}_{n\to\infty} A_n)^*$ is not a generally valid equation in the sense that the existence of either side implies that one of the other side also. (Cf. (5), p. 382.)

(vi) *If* $A_1, A_2, \cdots$ *is a regular sequence and* A *is normed, then* $(A_1, A_2, \cdots) \sim A$ *is equivalent to* $\lim_{n\to\infty} [[A_n - A]] = 0$.

PROOF: Ad (i): Choose an $(A_1, A_2, \cdots) \sim A$. Then $[[A]] = \lim_{n\to\infty} [[A_n]]$. Since all $[[A_n]] \geqq 0$, so $[[A]] \geqq 0$. And $[[A]] = 0$ means $\lim_{n\to\infty} [[A_n]] = 0$, that is, the equivalence of $A_1, A_2, \cdots$ and $0, 0, \cdots$. But as $(A_1, A_2, \cdots) \sim A$, $(0, 0, \cdots) \sim 0$, therefore this means $A = 0$.

Ad (ii), (iii), (iv), (v): Immediately by (ii), (iii), (iv), (iii) in Theorem III, respectively. (In the second part of (v) interchange A and B.)

Ad (vi): *Sufficiency*: Assume $(A_1, A_2, \cdots) \sim A$. Clearly $(-A_n, -A_n, \cdots) \sim -A_n$, hence by (ii) in Lemma 1.5.1. $(A_1 - A_n, A_2 - A_n, \cdots) \sim A - A_n$. Therefore $\lim_{m\to\infty} [[A_m - A_n]] = [[A - A_n]]$. Now $A_1, A_2, \cdots$ must be fundamental, so $\lim_{m,n\to\infty} [[A_m - A_n]] = 0$. Since $\lim_{m\to\infty} [[A_m - A_n]]$ exists, we have a fortiori $\lim_{n\to\infty} (\lim_{m\to\infty} [[A_m - A_n]]) = 0$, $\lim_{n\to\infty} [[A - A_n]] = 0$, that is $\lim_{n\to\infty} [[A_n - A]] = 0$.

Necessity: Assume $\lim_{n\to\infty} [[A_n - A]] = 0$. By (iii) above, $[[A_m - A_n]] = [[(A_m - A) - (A_n - A)]] \leqq [[A_m - A]] + [[A_n - A]]$. Hence the above assumption implies $\lim_{m,n\to\infty} [[A_m - A_n]] = 0$. So $A_1, A_2, \cdots$ being regular, is even fundamental. Therefore a normed A' with $(A_1, A_2, \cdots) \sim A'$ exists. Hence by our above result $\lim_{n\to\infty} [[A_n - A']] = 0$. Now again by (iii) above $[[A - A']] = [[(A_n - A') - (A_n - A)]] \leqq [[A_n - A']] + [[A_n - A]]$. As $\lim_{n\to\infty} [[A_n - A']] = \lim_{n\to\infty} [[A_n - A]] = 0$, this proves $[[A - A']] = 0$. Therefore by (i) above, $A - A' = 0$, $A = A'$. Hence $(A_1, A_2, \cdots) \sim A$.

LEMMA 1.5.2. *If* $(A_1, A_2, \cdots) \sim A$, $(B_1, B_2, \cdots) \sim B$, *then* $\lim_{n\to\infty} \overline{T}\mathfrak{M}(A_nB_n^*)$ *exists, and it depends on* A, B *only (not on the sequences* $A_1, A_2, \cdots$ *and* $B_1, B_2, \cdots$ *themselves!).*

PROOF: It suffices to prove this for $\lim_{n\to\infty} \Re\overline{T}\mathfrak{M}(A_nB_n^*)$ and $\lim_{n\to\infty} \Im\overline{T}\mathfrak{M}(A_nB_n^*)$ separately. And since the latter obtains from the former by replacing A and the A_n by iA and the iA_n, so we need to consider $\lim_{n\to\infty} \Re\overline{T}\mathfrak{M}(A_nB_n^*)$ only.

Now

$$\tfrac{1}{2}([[A_n + B_n]]^2 - [[A_n]]^2 - [[B_n]]^2)$$

$$= \tfrac{1}{2}(\overline{T}\mathfrak{M}[[A_n + B_n)(A_n + B_n)^*) - \overline{T}\mathfrak{M}(A_nA_n^*) - \overline{T}\mathfrak{M}(B_nB_n^*))$$

$$= \tfrac{1}{2}(\overline{T}\mathfrak{M}(A_nB_n^*) + \overline{T}\mathfrak{M}(B_nA_n^*))$$

$$= \tfrac{1}{2}(\overline{T}\mathfrak{M}(A_nB_n^*) + \overline{\overline{T}\mathfrak{M}(B_nA_n^*)}) = \Re\overline{T}\mathfrak{M}(A_nB_n^*),$$

$$\Re\,\overline{T}\mathfrak{M}(A_nB_n^*) = \tfrac{1}{2}([[A_n + B_n]]^2 - [[A_n]]^2 - [[B_n]]^2). \tag{33}$$

We have $(A_1, A_2, \cdots) \sim A$, $(B_1, B_2, \cdots) \sim B$, hence by (iii) in Lemma 1.5.1, also $(A_1 + B_1, A_2 + B_2, \cdots) \sim A + B$. So $\lim_{n\to\infty} [[A_n]] = [[A]]$, $\lim_{n\to\infty} [[B_n]] = [[B]]$, $\lim_{n\to\infty} [[A_n + B_n]] = [[A + B]]$. Consequently (33) gives

$$\lim_{n\to\infty} \Re\overline{T}\mathfrak{M}(A_nB_n^*) = \tfrac{1}{2}([[A + B]]^2 - [[A]]^2 - [[B]]^2).$$

This completes the proof.

DEFINITION 1.5.1. *Let A, B be normed. By Lemma* 1.5.2 $\lim_{n\to\infty} \overline{T}_{\mathfrak{M}}(A_nB_n^*)$ *exists for any two sequences $A_1, A_2, \cdots$ and $B_1, B_2, \cdots$ with $(A_1, A_2, \cdots) \sim A$, $(B_1, B_2, \cdots) \sim B$, and it depends on A, B only (not on the sequences $A_1, A_2, \cdots$ and $B_1, B_2, \cdots$ themselves!). We shall define the "inner product" of A, B, the number $\ll A, B \gg$, as this limit: $\ll A, B \gg = \lim_{n\to\infty} \overline{T}_{\mathfrak{M}}(A_nB_n^*)$.—*

Throughout what follows we shall make use of the results of Theorem IV and Definition 1.5.1, without referring to them explicitly.

THEOREM V. (i) *$\ll A, B \gg$ is linear in A, conjugate-linear in B, and of Hermitean symmetry in A, B. That is*

$$\ll aA, B \gg = a \ll A, B \gg,$$

$$\ll A' + A'', B \gg = \ll A', B \gg + \ll A'', B \gg,$$

$$\ll A, aB \gg = \bar{a} \ll A, B \gg,$$

$$\ll A, B' + B'' \gg = \ll A, B' \gg + \ll A, B'' \gg,$$

$$\ll A, B \gg = \overline{\ll B, A \gg}.$$

(ii) $$\ll A, A \gg = [[A]]^2.$$

(iii) $$\ll AD, B \gg = \ll A, BD^* \gg,$$

$$\ll DA, B \gg = \ll A, D^*B \gg,$$

where A, B are normed and $D \in \mathfrak{M}$.

PROOF: All these statements are immediate by Definition 1.5.1, and the properties of $\overline{T}_{\mathfrak{M}}(X)$ for X of finite rank. (Remember, in particular, that $\overline{T}_{\mathfrak{M}}(XY) = \overline{T}_{\mathfrak{M}}(YX)$ if either X or Y has a finite rank, and both are $\in \mathfrak{M}$.)—

We are now in a position to prove:

LEMMA 1.5.3. *A projection E is normed if and only if it is finite.*

PROOF: *Sufficiency*: If E is finite, then it is of finite rank, hence it is normed.

Necessity: Assume that E is normed. We want to prove that it is finite, that is: If $F \in \mathfrak{M}$ is a projection $\leqq E$, with $F \sim E$ ($\cdots$ $\mathfrak{M}$) ((**1**), p. 151, Definition 6.1.1), then $F = E$ ((**1**), p. 155, Definition 7.1.1). So we assume that F has the above properties (excepting of course $F = E$).

$F \sim E$($\cdots$ $\mathfrak{M}$) means $F = U^*U$, $E = UU^*$ for a partially isometric $U \in \mathfrak{M}$ ((**1**), p. 151, Definition 6.1.1), that is a U with $UU^*U = U$, $U^*UU^* = U^*$ ((**1**), p. 142, Lemma 4.3.2). Then $U = UU^*U = EU$ and $F = U^*U$ are both normed, along with E. $F \leqq E$ means $FE = F$.

Now we obtain, with the help of (i)–(iii) in Theorem V:

$$\ll E, F\gg = \ll E, F^*F\gg = \ll FE, F\gg = \ll F, F\gg = [[F]]^2,$$

$$\ll F, E\gg = \overline{\ll E, F\gg} = \overline{[[F]]^2} = [[F]]^2,$$

$$[[E-F]]^2 = \ll E-F, E-F\gg$$

$$= \ll E, E\gg + \ll F, F\gg - \ll E, F\gg - \ll F, E\gg = [[E]]^2 - [[F]]^2,$$

$$[[E-F]]^2 = [[E]]^2 - [[F]]^2. \tag{34}$$

Also:

$$[[E]]^2 = \ll E, E\gg = \ll UU^*, UU^*\gg = \ll UU^*U, U\gg = \ll U, U\gg,$$

$$[[F]]^2 = \ll F, F\gg = \ll U^*U, U^*U\gg = \ll UU^*U, U\gg = \ll U, U\gg,$$

$$[[E]]^2 = [[F]]^2. \tag{35}$$

Combining (34), (35) gives $[[E-F]]^2 = 0$, $[[E-F]] = 0$, and consequently $E = F$.

Thus the proof is completed.—

After this we could determine the normed A completely by investigating the spectral form of A^*A and the projections of its resolution of unity $E(\lambda)$, $-\infty < \lambda < +\infty$. We shall not do this here, however, because for our further purposes (especially for Lemma 2.2.1) Lemma 1.5.3 is sufficient.

1.6. We want to say a few words about the meaning of our above notions of *normedness* and *norm* in factors $\mathfrak{M}$ of various classes. The contents of this section, however, are not needed for the understanding of our subsequent deductions.

(A) $\mathfrak{M}$ *is of class* (I_n) ($n = 1, 2, \cdots$) or (II_1): $\mathfrak{M}$ is in this case of a finite class, hence every $\mathfrak{M}\eta\mathfrak{M}$ is finite, every $A \in \mathfrak{M}$ has finite rank. $\overline{T}_{\mathfrak{M}}(A)$ is defined for all $A \in \mathfrak{M}$ and coincides with the relative trace $T_{\mathfrak{M}}(A)$. (We can choose $\mathfrak{M} = \mathfrak{H}$ in Definitions 1.2.3, 1.2.4, and $D(\mathfrak{M})$ in the standard normalization.) Hence every A is normed, and $[[A]]$, $\ll A, B\gg$ mean the same thing as in (**2**), p. 241, Definition 4.3.1, and Lemma 4.3.2.

(B) $\mathfrak{M}$ *is of class* (I_∞): We might as well take $\mathfrak{M}$ as the ring $\mathfrak{B}$ of all bounded operators in $\mathfrak{H}$. (Cf. (**1**), p. 173, Lemma 8.6.1.) Let $\varphi_1, \varphi_2, \cdots$ be a complete, normalized, orthogonal set in $\mathfrak{H}$ and represent every $A \in \mathfrak{M} = \mathfrak{B}$ by its *matrix*:

$$A \approx \{a_{ij}\}, \qquad i, j = 1, 2, \cdots,$$

$$a_{ij} = (A\varphi_i, \varphi_j).$$

Choose $D(\mathfrak{M})$ in the standard normalizations; then it is the common notion of dimension. (Cf. loc. cit. above.) Hence Rank (A) has its usual meaning too. We show now:

B.1. (i) *If* Rank (A) *is finite, then* $\sum_1^\infty a_{ii}$ *is absolutely convergent, and* $\overline{T}_{\mathfrak{B}}(A) = \sum_1^\infty a_{ii}$.

(ii) *If* Rank (A) *is finite, then* $\sum_{i,j=1}^\infty |a_{ij}|^2$ *is absolutely convergent, and* $[[A]]^2 = \sum_{i,j=1}^\infty |a_{ij}|^2$.

Proof: Ad (i): If Rank (A) is finite, then Rank $(A) = D$ ([Range A]) $= k = 0, 1, 2, \cdots$, hence [Range A] $= [\psi_1, \cdots, \psi_k]$ where $\psi_1, \cdots, \psi_k$ are a normalized orthogonal set. Put $E_l = P_{[\psi_l]}$, $l = 1, \cdots, k$, then $P_{[\text{Range } A]} = E_1 + \cdots + E_k$, so $A = P_{[\text{Range } A]}A = E_1A + \cdots + E_kA$. Each E_lA has at most rank 1 so it suffices to prove (i) when the rank is at most 1. In this case clearly $A = \{a_{ij}\}$, $a_{ij} = \alpha_i\beta_j$. If all $\alpha_i = 0$ or all $\beta_j = 0$, then all $a_{ij} = 0$ in which case our statement is trivially true. Hence we may assume that some $\alpha_i \neq 0$ and some $\beta_j \neq 0$. As every $\sum_{j=1}^\infty |a_{ij}|^2$ and every $\sum_{i=1}^\infty |a_{ij}|^2$ must be finite,[8] this implies that both $\sum_{i=1}^\infty |\alpha_i|^2$ and $\sum_{j=1}^\infty |\beta_j|^2$ are finite. So we have:

$$0 < \sum_{i=1}^\infty |\alpha_i|^2, \sum_{j=1}^\infty |\beta_j|^2 < +\infty.$$

This proves in the first place that $\sum_{i=1}^\infty a_{ii} = \sum_{i=1}^\infty \alpha_i\beta_i$ is absolutely convergent (by Schwarz's well-known inequality). Now put $B = \{b_{ij}\}$, $C = \{c_{ij}\}$, where

$$b_{ij}\begin{cases} = \alpha_j & \text{for } j = 1 \\ = 0 & \text{otherwise} \end{cases}, \qquad c_{ij}\begin{cases} = \beta_j & \text{for } i = 1 \\ = 0 & \text{otherwise} \end{cases}.$$

Then $BC = A$, $CB = \{d_{ij}\}$, where

$$d_{ij}\begin{cases} = \sum_{i=1}^\infty \alpha_i\beta_i & \text{for } i, j = 1 \\ = 0 & \text{otherwise} \end{cases}.$$

So $CB = (\sum_{i=1}^\infty \alpha_i\beta_i)E_1$, where $E_1 = P_{[\varphi_1]}$. B, C are clearly of finite rank $(= 1)$. Now we have (remember $D(E_1) = 1$, hence $\overline{T}_{\mathfrak{M}}(E_1) = 1$):

$$\overline{T}_{\mathfrak{M}}(A) = \overline{T}_{\mathfrak{M}}(BC) = \overline{T}_{\mathfrak{M}}(CB) = \left(\sum_{i=1}^\infty \alpha_i\beta_i\right)\overline{T}_{\mathfrak{M}}(E_1) = \sum_{i=1}^\infty \alpha_i\beta_i.$$

Thus the proof is completed.

Ad (ii): Immediately by (i), since $A^*A = \{g_{ij}\}$, $g_{ij} = \sum_{k=1}^\infty \overline{a_{ki}}a_{kj}$, hence $g_{ii} = \sum_{k=1}^\infty |a_{ki}|^2$, $\sum_{i=1}^\infty g_{ii} = \sum_{k,i=1}^\infty |a_{ki}|^2$.

B.2. *A is normed if and only if* $\sum_{i,j=1}^\infty |a_{ij}|^2$ *is finite. Again* $[[A]]^2 = \sum_{i,j=1}^\infty |a_{ij}|^2$.

---

[8] $\sum_{j=1}^\infty |a_{ij}|^2 = \sum_{j=1}^\infty |(A\varphi_i, \varphi_j)|^2 = ||A\varphi_i||^2$, so $\sum_{j=1}^\infty |a_{ij}|^2$ is finite. The same results for $\sum_{j=1}^\infty |a_{ij}|^2$ by considering $A^* = \{\overline{a_{ji}}\}$ in place of $A = \{a_{ij}\}$.

PROOF: *Necessity*: Assume that A is normed. Put $E_n = P_{[\varphi_1, \cdots, \varphi_n]}$. Then $|||E_n||| = 1$, hence E_nAE_n is normed, and $[[E_nAE_n]] \leqq [[A]]$. Now E_nAE_n has even finite rank, along with E_n, and $A \approx \{a_{ij}\}$ implies $E_nAE_n \approx \{a_{ij}^{(n)}\}$, where

$$a_{ij}^{(n)} \begin{cases} = a_{ij} & \text{for } i, j = 1, \cdots, n \\ = 0 & \text{otherwise} \end{cases}.$$

Hence by (ii) in B.1:

$$[[E_nAE_n]]^2 = \sum_{i,j=1}^{\infty} |a_{ij}^{(n)}|^2 = \sum_{i,j=1}^{n} |a_{ij}|^2,$$

and consequently $\sum_{i,j=1}^{n} |a_{ij}|^2 \leqq [[A]]^2$ for every $n = 1, 2, \cdots$. This proves the finiteness of $\sum_{i,j=1}^{\infty} |a_{ij}|^2$.

Sufficiency, and value of $[[A]]^2$: Assume that $\sum_{i,j=1}^{\infty} |a_{ij}|^2$ is finite. Form E_n as above, put $A_n = E_nAE_n$. Each A_n has finite rank, and $|||A_n||| = |||E_nAE_n||| \leqq |||A|||$. Hence the sequence $A_1, A_2, \cdots$ is regular. It is also fundamental: If $m \geqq n$, then $A_m - A_n \approx \{a_{ij}^{(m)} - a_{ij}^{(n)}\}$, and

$$a_{ij}^{(m)} - a_{ij}^{(n)} \begin{cases} = a_{ij} & \text{for } i, j = 1, \cdots, m, \text{ but not } i, j = 1, \cdots, n \\ = 0 & \text{otherwise} \end{cases},$$

and consequently (using (ii) in B.1)

$$[[A_m - A_n]]^2 = \sum_{i,j=1}^{\infty} |a_{ij}^{(m)} - a_{ij}^{(n)}|^2 = \sum_{i,j=1}^{m} |a_{ij}|^2 - \sum_{i,j=1}^{n} |a_{ij}|^2.$$

So for $m \gtrless n$

$$[[A_m - A_n]]^2 = \left| \sum_{i,j=1}^{m} |a_{ij}|^2 - \sum_{i,j=1}^{n} |a_{ij}|^2 \right|,$$

and so the finiteness of $\sum_{i,j=1}^{\infty} |a_{ij}|^2$ implies $\lim_{m,n\to\infty} [[A_m - A_n]] = 0$.

Finally $\lim \operatorname{str}_{n\to\infty} E_n = I$, hence $\lim \operatorname{str}_{n\to\infty} A_n = \lim \operatorname{str}_{n\to\infty} E_nAE_n = A$. Thus we have

$$(A_1, A_2, \cdots) \sim A.$$

Therefore A is normed, and (using (ii) in B.1)

$$[[A]]^2 = \lim_{n\to\infty} [[A_n]]^2 = \lim_{n\to\infty} \sum_{i,j=1}^{\infty} |a_{ij}^{(n)}|^2 = \lim_{n\to\infty} \sum_{i,j=1}^{n} |a_{ij}|^2 = \sum_{i,j=1}^{\infty} |a_{ij}|^2.$$

Thus the proof is completed.—

B.2 shows that in this case the class of all normed A coincides with the well-known class of E. Schmidt.

We finally observe that in this case the condition $\lim_{m,n\to\infty} [[A_m - A_n]] = 0$ of Definition 1.4.2 (which implies the existence of $\lim_{n\to\infty} [[A_n]]$ hence the boundedness of the sequence $[[A_1]], [[A_2]], \cdots$) implies (iii) in Definition 1.4.1. So in the present case (I_∞) condition (iii) in Definition 1.4.1 is superfluous—which is not true in the case (II_∞), as will be seen in C.1 below.

Indeed we have:

B.3 $|||A||| \leqq [[A]]$ *if* A *is normed.*

PROOF: That is: $||Af|| \leqq [[A]] \cdot ||f||$ for every $f \in \mathfrak{H}$. We may obviously assume $||f|| = 1$, and then we can choose the complete, normalized orthogonal set $\varphi_1, \varphi_2, \cdots$ so that $\varphi_1 = f$. So we want to prove $||A\varphi_1|| \leqq [[A]]$ that is, $||A\varphi_1||^2 \leqq [[A]]^2$. This, however, is obvious.[9]

(C) $\mathfrak{M}$ *is of class* (II_∞): In this case our notions are genuinely new. We only wish to observe that the opposite of B.3 is true, which justifies our formulation of Definitions 1.4.1, 1.4.2. (Cf. the above discussion before B.3.) Indeed we have:

C.1. *Even for an* A *of finite rank* $|||A|||/[[A]]$ *can be arbitrarily great.*

PROOF: Given any $C > 0$, choose a projection $E \in \mathfrak{M}$ with $D(E) > 0, \leqq 1/C^2$. Then E is finite, hence of finite rank. $|||E||| = 1$, $[[E]] = (D(E))^{\frac{1}{2}} \leqq 1/C$, and consequently $|||E|||/[[E]] \geqq C$.

(D) $\mathfrak{M}$ *is of class* (III_∞): If A is of finite rank, then [Range A] is finite, hence (since $\mathfrak{M}$ is of class (III_∞)!) [Range A] $= 0$, $A = 0$. Hence for every regular, and a fortiori for every fundamental sequence $A_1, A_2, \cdots$ necessarily $A_1 = A_2 = \cdots = 0$. Thus the only normed A is $A = 0$.

In this case, therefore, all the notions we have defined are vacuous.—

In spite of this, just these notions will help us to prove that the examples of factors which we shall construct in §§4.3–4.4, belong in certain cases to class (III_∞). (Cf. in particular §4.3, Theorem IX.) The procedure will be essentially the reverse of (C) above: We shall succeed in showing that in those factors $\mathfrak{M}$ every finite projection E must be $= 0$, the proof being mainly based on the properties of the notion of normedness. From this we then conclude that such an $\mathfrak{M}$ must be of class (III_∞). Any direct attempt to prove such a thing will show that of all possibilities alternative to (III_∞) the case (II_∞) is the hardest one to exclude. It is therefore very appropriate that our notions of normedness and norm are in this case genuinely new.

We also want to point out, that Theorems IV, V show, that the normed operators in $\mathfrak{M}$ form a (possibly) incomplete Hilbert space. This space is actually complete in the following cases: (I_n): Obvious. (I_∞): Clear by (B) above. (III_∞): Clear, since it contains 0 only, cf. (D) above. In the cases (II_1), (II_∞) it is not complete, but can be completed to a space $Q(\mathfrak{M})$, by adjoining to $\mathfrak{M}$ certain (but not all) unbounded operators $\eta \mathfrak{M}$. This process was carried out in detail for (II_1) in (**2**), pp. 236–242. In this connection C.1 above is significant. We will not go here any further into this subject.

CHAPTER II. MAXIMUM ABELIAN RINGS AND THE OPERATIONS $A^{|E_1|E_2|\cdots}$

2.1. DEFINITION 2.1.1. *For any bounded operator* A *and any projection* E *we define*

$$A^{|E} = EAE + (I - E)A(I - E).$$

[9] $||A\varphi_1||^2 = \sum_{j=1}^{\infty} |a_{1j}|^2$ by footnote 8, $[[A]]^2 = \sum_{i,j=1}^{\infty} |a_{ij}|^2$ by B. 2.

Consequently

$$A - A^{|E} = (I - E)AE + EA(I - E).$$

LEMMA 2.1.1. (i) *If* $A, E \in \mathfrak{M}$ *then* $A^{|E} \in \mathfrak{M}$.
(ii) *If* $A, E \in \mathfrak{M}$ *and* A *is of finite rank, then* $A^{|E}$ *is also of finite rank.*

PROOF: Both statements are obvious.

LEMMA 2.1.2. $A^{|E}$ *commutes with* E.

PROOF: Results by an obvious computation.

LEMMA 2.1.3. *If* $E_1, \cdots, E_n$ *commute with each other, and if* $(\pi_1, \cdots, \pi_n)$ *is a permutation of* $(1, \cdots, n)$, *then*

$$A^{|E_1|\cdots|E_n} = A^{|E_{\pi_1}|\cdots|E_{\pi_n}}.$$

PROOF: It is obviously sufficient to consider the case where $(\pi_1, \cdots, \pi_n)$ is just a transposition of two neighbors, say of m, $m+1$ $(m = 1, \cdots, n-1)$. Then we must prove $B^{|E_m|E_{m+1}} = B^{|E_{m+1}|E_m}$, put $B = A^{|E_1\cdots|E_{m-1}}$, and apply $^{|E_{m+2}|\cdots|E_n}$ to the result. Writing A, E, F for B, E_m, E_{m+1} we see: We have to show

$$A^{|E|F} = A^{|F|E}$$

if E, F commute. This, however, results by an obvious computation.

COROLLARY. *If* $E_1, \cdots, E_n$ *commute with each other, then* $A^{|E_1|\cdots|E_n}$ *commutes with all of them.*

PROOF: Consider an $i = 1, \cdots, n$. Let $(\pi_1, \cdots, \pi_n)$ be a permutation of $(1, \cdots, n)$ with $\pi_n = i$. Then $A^{|E_1|\cdots|E_n} = A^{|E_{\pi_1}|\cdots|E_{\pi_n}}$ by Lemma 2.1.3, and this commutes with $E_{\pi_n} = E_i$ by Lemma 2.1.2.

LEMMA 2.1.4. $|||\, A^{|E_1|\cdots|E_n} \,||| \leqq |||\, A \,|||$.

PROOF: We shall prove $|||\, A^{|E_1|\cdots|E_{n-1}|E_m} \,||| \leqq |||\, A^{|E_1|\cdots|E_{m-1}} \,|||$. From this the desired result obtains by putting $m = 1, \cdots, n$. Write in the above inequality A, E for $A^{|E_1|\cdots|E_{m-1}}, E_m$. Then we obtain $|||\, A^{|E} \,||| \leqq |||\, A \,|||$. In other words: We want to prove

$$|\,(A^{|E}f, g)\,| \leqq |||\, A \,|||\cdot\|f\|\cdot\|g\|. \tag{36}$$

Now clearly

$$(A^{|E}f, g) = (EAEf, g) + ((I - E)A(I - E)f, g)$$
$$= (AEf, Eg) + (A(I - E)f, (I - E)g).$$

Hence

$$|\,(A^{|E}f, g)\,| \leqq |||\, A \,|||\cdot(\|Ef\|\cdot\|Eg\| + \|(I - E)f\|\cdot\|(I - E)g\|). \tag{37}$$

By Schwarz's well-known inequality

$$\|Ef\|\cdot\|Eg\| + \|(I - E)f\|\cdot\|(I - E)g\|$$
$$\leqq [(\|Ef\|^2 + \|(I - E)f\|^2)(\|Eg\|^2 + \|(I - E)g\|^2)]^{\frac{1}{2}}.$$

But

$$\| Ef \|^2 + \| (I - E)f \|^2 = (Ef, Ef) + ((I - E)f, (I - E)f)$$
$$= (Ef, f) + ((I - E)f, f) = (f, f) = \| f \|^2,$$

similarly

$$\| Eg \|^2 + \| (I - E)g \|^2 = \| g \|^2.$$

Therefore

$$\| Ef \| \cdot \| Eg \| + \| (I - E)f \| \cdot \| (I - E)g \| \leqq \| f \| \cdot \| g \|.$$

And this carries (37) over into (36), and thereby completes the proof.

LEMMA 2.1.5. *Assume that* A, E_1, E_2, $\cdots \in \mathfrak{M}$, A *of finite rank, the* E_1, E_2, $\cdots$ *commute with each other. Then*

$$[[A^{|E_1|\cdots|E_m} - A^{|E_1|\cdots|E_n}]]^2 = [[A^{|E_1|\cdots|E_m}]]^2 - [[A^{|E_1|\cdots|E_n}]]^2$$

if $m \leqq n$, $m, n = 1, 2, \cdots$.

PROOF: All $A^{|E_1|E_2|\cdots}$ are of finite rank by Lemma 2.1.1, (ii), so all quantities which occur in the above equation are defined. We want to prove

$$[[X + Y]]^2 = [[X]]^2 + [[Y]]^2 \tag{38}$$

for $X = A^{|E_1|\cdots|E_n}$, $Y = A^{|E_1|\cdots|E_m} - A^{|E_1|\cdots|E_n}$.

Now

$$[[X + Y]]^2 = \ll X + Y, X + Y \gg$$
$$= \ll X, X \gg + \ll Y, Y \gg + \ll X, Y \gg + \ll Y, X \gg$$
$$= [[X]]^2 + [[Y]]^2 + 2\Re \ll X, Y \gg,$$

hence (38) is certainly a consequence of

$$\ll X, Y \gg = 0. \tag{39}$$

We shall prove (39) for the above X, Y. And since

$$A^{|E_1|\cdots|E_m} - A^{|E_1|\cdots|E_n} = \sum_{p=m}^{n-1} (A^{|E_1|\cdots|E_p} - A^{|E_1|\cdots|E_p|E_{p+1}}),$$

so we may prove (39) for $X = A^{|E_1|\cdots|E_n}$, $Y = A^{|E_1|\cdots|E_p} - A^{|E_1|\cdots|E_p|E_{p+1}}$ with $p < n$, $p + 1 \leqq n$.

Let $(\pi_1, \cdots, \pi_n)$ be a permutation of $(1, \cdots, n)$ with $\pi_n = p + 1$. Then Lemma 2.1.3 gives $A^{|E_1|\cdots|E_n} = A^{|E_{\pi_1}|\cdots|E_{\pi_{n-1}}|E_{\pi_n}} = A^{|E_{\pi_1}|\cdots|E_{\pi_{n-1}}|E_{p+1}}$. Write B, C, E for $A^{|E_{\pi_1}|\cdots|E_{\pi_{n-1}}}$, $A^{|E_1|\cdots|E_p}$, E_{p+1}, then we see: We need to prove (39) for $X = B^{|E}$, $Y = C - C^{|E}$ only. By Definition 2.1.1, $B^{|E} = EBE + (I - E)B(I - E)$, $C - C^{|E} = (I - E)CE + EC(I - E)$. Hence we need to prove (39) only for the following combinations of X, Y:

$$\begin{cases} X = EBE & \text{or} \quad (I - E)B(I - E), \\ Y = (I - E)CE & \text{or} \quad EC(I - E). \end{cases} \tag{40}$$

Now we have for the various alternatives of (40)

$$(41)\qquad \begin{cases}\text{Either} & FX = X, FY = 0 \quad \text{for} \quad F = E \quad \text{or} \quad I - E,\\ \text{or} & XF = X, YF = 0 \quad \text{for} \quad F = F \quad \text{or} \quad I - E.\end{cases}$$

Owing to (41) we now have:

$$\begin{aligned}&\text{Either} \quad \ll X, Y\gg = \ll FX, Y\gg = \ll X, FY\gg = 0,\\ &\text{or} \quad \ll X, Y\gg = \ll XF, Y\gg = \ll X, YF\gg = 0.\end{aligned}$$

Hence (39) holds in any case.

Thus the proof is completed.

LEMMA 2.1.6. *Let $A, E_1, E_2, \cdots$ be as in Lemma* 2.1.5. *Put $A_n = A^{|E_1|\cdots|E_n}$ for $n = 1, 2, \cdots$. Then $A_1, A_2, \cdots$ is a fundamental sequence.*

PROOF: $A_1, A_2, \cdots$ is a regular sequence: We must verify the conditions (i)–(iii) in Definition 1.4.1.

Ad (i): Obvious.

Ad (ii): Immediate by repeated application of Lemma 2.1.1, (ii).

Ad (iii): Immediate by Lemma 2.1.4.

$A_1, A_2, \cdots$ is a fundamental sequence: We must verify further the condition of Definition 1.4.2, that is

$$(42)\qquad \lim_{m,n\to\infty} [[A_m - A_n]]^2 = 0.$$

Now Lemma 2.1.5 gives $[[A_m]]^2 \geqq [[A_n]]^2$ for $m \leqq n$,

$$[[A_1]]^2 \geqq [[A_2]]^2 \geqq \cdots \geqq 0.$$

Therefore $\lim_{n\to\infty} [[A_n]]^2$ exists. Hence

$$(43)\qquad \lim_{m,n\to\infty} |\,[[A_m]]^2 - [[A_n]]^2\,| = 0.$$

But

$$(44)\qquad [[A_m - A_n]]^2 = |\,[[A_m]]^2 - [[A_n]]^2\,|.$$

Indeed: By symmetry in m, n we need only to prove (44) for $m \leqq n$, and then it follows from Lemma 2.1.5.

Now (43), (44) give together (42).

DEFINITION 2.1.2. *Let $A, E_1, E_2, \cdots$ be as in Lemma* 2.1.5. *Then there exists by Lemma* 2.1.6 *a unique $\bar{A}$, so that*

$$(A_1, A_2, \cdots) \sim \bar{A}, \quad \textit{where} \quad A_n = A^{|E_1|\cdots|E_n}.$$

We shall write

$$\bar{A} = A^{|E_1|E_2|\cdots}.$$

THEOREM VI. *Let $A, E_1, E_2, \cdots$ be as in Lemma* 2.1.5. *Then we have:*

(i) $A^{|E_1|E_2|\cdots}$ *is normed.*

(ii) $A^{|E_1|E_2|\cdots}$ *commutes with $E_1, E_2, \cdots$.*

PROOF: Ad (i): Obvious by Definition 2.1.2.
Ad (ii): Consider a fixed $i = 1, 2, \cdots$. For $n \geqq i$ $A_n = A^{|E_1|\cdots|E_n}$ commutes with E_i by the Corollary to Lemma 2.1.3. But $A^{|E_1|E_2|\cdots} = \lim \operatorname{str}_{n\to\infty} A_n$ by Definition 2.1.2, hence $A^{|E_1|E_2|\cdots}$ commutes also with E_i.

2.2. DEFINITION 2.2.1. *A ring $\mathfrak{L} \subset \mathfrak{M}$ will be said to be "purely infinite" if every projection $E \in \mathfrak{L}$ is either $= 0$ or infinite (with respect to $\mathfrak{M}$, cf.* (1), *p.* 155, *Definition* 7.1.1).—

Observe that the following terms: finite, infinite (cf. loc. cit. above) of finite rank (Definition 1.2.1) normed (Definition 1.4.6) are always to be understood with respect to $\mathfrak{M}$. Being purely infinite, however, is defined in Definition 2.2.1 above for any ring $\mathfrak{L} \subset \mathfrak{M}$. But it has, of course, for $\mathfrak{L} = \mathfrak{M}$ the old meaning of $\mathfrak{M}$ being purely infinite ((1), p. 172, Theorem VIII): That $\mathfrak{M}$ belongs to class (III_∞).

We prove now:

LEMMA 2.2.1. *A ring $\mathfrak{L} \subset \mathfrak{M}$ is purely infinite, if and only if the only normed $A \in \mathfrak{L}$ is $A = 0$.*

PROOF: *Sufficiency*: Assume that the only normed $A \in \mathfrak{L}$ is $A = 0$. Apply this to projections $E \in \mathfrak{L}$. They must be $= 0$ or infinite. Hence $\mathfrak{L}$ is purely infinite.

Necessity: Assume that $\mathfrak{L}$ is purely infinite. Consider a normed $A \in \mathfrak{L}$. Then A^*A is also normed and also $\in \mathfrak{L}$ along with A.

We now proceed as in the proof of Lemma 1.3.2, and obtain the spectral form of A^*A, in particular the equations

$$(A^*Af, g) = \int_{-\infty}^{\infty} \lambda \, d(E(\lambda)f, g), \tag{6}$$

$$\| Af \|^2 = \int_{-\infty}^{\infty} \lambda \, d(\| E(\lambda)f \|^2), \tag{7}$$

$$\lambda < 0 \quad \text{implies} \quad E(\lambda) = 0 \tag{8}$$

where $E(\lambda)$, $-\infty < \lambda < +\infty$, is a *resolution of unity*. (Cf. loc. cit. above.)

Consider now a $\lambda_0 > 0$. Form the bounded function

$$\alpha(\lambda) \begin{cases} = \dfrac{1}{\lambda} & \text{for} \quad \lambda \geqq \lambda_0 \\ = 0 & \text{otherwise} \end{cases}.$$

We can form the operator $B = \alpha(A^*A)$ which is characterized by

$$(Bf, g) = \int_{-\infty}^{\infty} \alpha(\lambda) \, d(E(\lambda)f, g) \tag{45}$$

((6), pp. 202–205), and which belongs to $\mathfrak{L}$ along with A^*A ((6), p. 213, Theorem 6—only the easy, sufficiency part of this theorem is needed). Now $B \cdot A^*A = \beta(A^*A)$, where

$$\beta(\lambda) = \lambda\alpha(\lambda)\begin{cases} = 1 & \text{for } \lambda \geqq \lambda_0 \\ = 0 & \text{otherwise} \end{cases}.$$

((6), pp. 205–206, property e) on p. 205); hence

$$(B \cdot A^*Af, g) = \int_{-\infty}^{\infty} \beta(\lambda)\, d(E(\lambda)f, g) = ((I - E(\lambda_0))f, g)$$

((6), pp. 202–205), and so

$$B \cdot A^*A = I - E(\lambda_0).$$

B, A^*A belong both to $\mathfrak{L}$ and A^*A is normed; hence $B \cdot A^*A = I - E(\lambda_0)$ belongs also to $\mathfrak{L}$ and is also normed. So we have shown:

$$\lambda > 0 \quad \text{implies} \quad E(\lambda) = I. \tag{46}$$

(8), (46) and (7) give together $\| Af \|^2 = 0$, $Af = 0$ for all $f \in \mathfrak{H}$. Hence $A = 0$.

This completes the proof.

LEMMA 2.2.2. *Consider an Abelian ring $\mathfrak{A} \subset \mathfrak{M}$. There exists a sequence of projections $E_1, E_2, \cdots \in \mathfrak{A}$ with $\mathfrak{A} = \mathfrak{R}(E_1, E_2, \cdots)$. (Cf.* (5), 401, *lower part of the page.) Since $\mathfrak{A}$ is Abelian, these $E_1, E_2, \cdots$ commute with each other. So we can form $A^{|E_1|E_2|\cdots}$ for every A of finite rank (Definition* 2.1.2).[10] *We consider $\mathfrak{A}$ and $E_1, E_2, \cdots$ as given.*

We make the following assumptions:

(i) *The ring $\mathfrak{M} \cdot \mathfrak{A}' (\subset \mathfrak{M})$ is purely infinite.*

(ii) *For every finite projection $F (\in \mathfrak{M})$*

$$F^{|E_1|E_2|\cdots} = 0 \quad \textit{implies} \quad F = 0.$$

Then $\mathfrak{M}$ is of class (III_∞).

PROOF: Let $F (\in \mathfrak{M})$ be a finite projection. Then we can form $F^{|E_1|E_2|\cdots}$. $F^{|E_1|E_2|\cdots}$ is normed ((i) in Theorem VI), and so in particular $\in \mathfrak{M}$. Furthermore $F^{|E_1|E_2|\cdots}$ commutes with $E_1, E_2, \cdots$ ((ii) in Theorem VI), hence with all of $\mathfrak{A} = \mathfrak{R}(E_1, E_2, \cdots)$ and so it is $\in \mathfrak{A}'$. Thus $F^{|E_1|E_2|\cdots} \in \mathfrak{M} \cdot \mathfrak{A}'$. Now $F^{|E_1|E_2|\cdots}$ is normed and $\mathfrak{M} \cdot \mathfrak{A}'$ is purely infinite by (i), hence Lemma 2.2.1 gives $F^{|E_1|E_2|\cdots} = 0$. And (ii) permits us therefore to conclude that $F = 0$.

In other words: A projection $F \in \mathfrak{M}$ is either $= 0$ or infinite. Hence $\mathfrak{M}$ is of class (III_∞).—

The above Lemma 2.2.2 is the device by which we are going to establish

[10] We could show that $A^{|E_1|E_2|\cdots}$ depends only on A and $\mathfrak{A}$ (and not on the sequence $E_1, E_2, \cdots$ itself). We omit the proof, because this fact is immaterial for our further purposes.

that some of our examples $\mathfrak{M}$ are of class (III_∞). (Cf. Lemma 4.3.5.) We shall apply it to rings $\mathfrak{A}$ which are $= \mathfrak{M} \cdot \mathfrak{A}'$, so-called *maximum Abelian rings in* $\mathfrak{M}$. The general theory of such rings is of some interest, but we do not propose to go into it in this connection.

Chapter III. Construction of the Rings $\mathfrak{M}$, $\mathfrak{M}'$

3.1. This chapter will be devoted to the construction and investigation of a family of examples of factors $\mathfrak{M}$ in appropriate Hilbert spaces $\mathfrak{H}$. Some of these $\mathfrak{M}$ will be shown to be of class (III_∞). More precisely: We shall succeed in determining the class of each $\mathfrak{M}$ of this family, and we shall see that all classes (I_n)–(III_∞) occur among them.

The examples which we shall consider are a generalization of those of (1), pp. 192–209, which belonged to the classes (I_n)–(II_∞). We desire, however, to treat the intervening notion of (generalized) Lebesgue measure in a slightly different way than was done loc. cit. above. Furthermore, this part of our discussion will have to be made on a broader basis than was necessary there. For these reasons we begin with a discussion of the properties of (generalized) Lebesgue measure and integration in an arbitrary space S.

3.2. Definition 3.2.1. *Let S be a fixed set. A system Γ of subsets of S will be called a "Borel-system" if it has the following properties:*

(i) Θ (*the empty set*) $\epsilon \Gamma$.
(ii) $M \epsilon \Gamma$ *implies* $S - M \epsilon \Gamma$.
(iii) $M_1, M_2, \cdots \epsilon \Gamma$ *imply* $M_1 + M_2 + \cdots \epsilon \Gamma$.

Lemma 3.2.1. *For a Borel system Γ in S we have further:*

(i) $S \epsilon \Gamma$.
(ii) $M, N \epsilon \Gamma$ *imply* $M - N \epsilon \Gamma$.
(iii) $M_1, M_2, \cdots \epsilon \Gamma$ *imply* $M_1 \cdot M_2 \cdot \cdots \epsilon \Gamma$.

Proof: We prove these statements in a somewhat changed order:

Ad (i): By (i), (ii) in Definition 3.2.1, since $S = S - \Theta$.
Ad (iii): By (ii), (iii) in Definition 3.2.1, since

$$M_1 \cdot M_2 \cdot \cdots = S - ((S - M_1) + (S - M_2) + \cdots).$$

Ad (ii): By (ii) Definition 3.2.1 and (iii) above, since $M - N = M \cdot (S - N)$.

Definition 3.2.2. *Let Δ be an arbitrary system of subsets of S. Then there exists a minimum Borel-system $\Gamma \supset \Delta$ in S,*[11] *which will be denoted by $\Gamma(\Delta)$.*

[11] The intersection of all Borel-systems $\Gamma \supset \Delta$ in S.

DEFINITION 3.2.3. *A Borel-system* Γ *in* S *will be called "separable," if there exists a system* Δ *of subsets of* S *with the following properties*:

(i) $\Gamma = \Gamma(\Delta)$.
(ii) *If* $x, y \in S$ *and if* $x \in M$ *is equivalent to* $y \in M$ *for all* $M \in \Delta$, *then* $x = y$.
(iii) Δ *is finite or enumerably infinite.*—

We can now define our notion of (generalized) Lebesgue measure. (To be called Γ-*measure*, cf. below.)

DEFINITION 3.2.4. *Let a separable Borel-system* Γ *be given in* S. *A function* $\mu(M)$ *will be called a "*Γ-*measure" if it possesses the following properties*:

(i) $\mu(M)$ *is defined for all* $M \in \Gamma$ *and only those.*
(ii) *The values of* $\mu(M)$ *are real numbers* $\geqq 0, \leqq +\infty$.
(iii) *There exists a sequence* $T^{(1)}, T^{(2)}, \cdots \in \Gamma$ *with* $T^{(1)} + T^{(2)} + \cdots = S$, *such that all* $\mu(T^{(i)}) < +\infty$.
(iv) *If* $M_1, M_2, \cdots \in \Gamma$ *and* $i \neq j$ *implies* $M_i \cdot M_j = \Theta$ *then*

$$\mu(M_1 + M_2 + \cdots) = \mu(M_1) + \mu(M_2) + \cdots .$$

We shall always consider measures with $\mu(S) \neq 0$.

DEFINITION 3.2.5. *A complex valued function* $f(x)$, *defined for all* $x \in S$, *will be called "*Γ-*measurable," if the sets*

$$(x; \Re f(x) > a), \qquad (x; \Im f(x) > a)$$

are $\in \Gamma$ *for every real* a.—

It is not necessary to develop the properties of the class of all Γ-measurable functions in detail since they are the same and obtained in precisely the same manner as those of the class of all Lebesgue-measurable functions of one real variable. (Cf. any standard exposition of that subject. For a systematic discussion of such function classes as ours, cf. e.g. **(0)**, pp. 232–247. Our Γ-measurable functions $f(x)$ are to be characterized in the terminology loc. cit. as follows: $\Re f(x)$, $\Im f(x)$ are both functions of class $(\Gamma, *)$ or, which is equivalent in this case, of class (Γ, Γ).) It may suffice to observe that the following functions in particular are Γ-measurable:

(i) The *characteristic function* $\theta_M(x) \begin{cases} = 1 \text{ for } x \in M \\ = 0 \text{ otherwise} \end{cases}$ of any $M \in \Gamma$.
(ii) $af(x)$, $\overline{f(x)}$ along with $f(x)$, a being any complex constant.
(iii) $f(x) + g(x)$, $f(x)g(x)$, Max $(f(x), g(x))$, Min $(f(x), g(x))$ along with $f(x), g(x)$.
(iv) $f(x) = \lim_{n\to\infty} f_n(x)$ (if this limit exists for all $x \in S$) along with $f_1(x)$, $f_2(x), \cdots$.

For the Γ-measurable functions $f(x)$ we now define the notion of Γ, μ-*integrability* (with respect to the Γ-measure μ!) and the numerical value of the Γ, μ-*integral*

$$\int_M f(x)\, d_\mu x \qquad (M \in \Gamma) \tag{47}$$

in the usual way. (Cf. any standard exposition of Lebesgue-integration. The observations of (**1**), p. 194, top of that page, apply to the present situation.) There is again no need to enumerate the well-known details. Let us mention this, however: For a general, complex valued, $f(x)$ we restrict, as usual, (47) to the case where $f(x)$ is even Γ, μ-summable over M that is, where $\int_M |f(x)|\, d_\mu x$ exists and is finite. In such situations, however, where $f(x)$ is restricted to real and $\geqq 0$ functions, we shall consider (47) as meaningful for all Γ-measurable $f(x)$, expressly permitting in this case the value $+\infty$ also for the integral. This convention too agrees with the general usage.

Throughout the rest of this section, as well as in §§3.3–3.7 which follow, we assume that S, Γ μ are given, and also that Γ is a separable Borel-system. Other Γ-measures ν, ν', ν'' and groups $\mathfrak{G}$, which will be considered, may vary.

We construct next a functional Hilbert space over S.

DEFINITION 3.2.6. *Let $\mathfrak{H}_S = \mathfrak{H}_{S,\Gamma,\mu}$ be the set of all (complex) valued Γ-measurable functions $f(x)$ $(x \in S)$ with a finite $\int_S |f(x)|^2\, d_\mu x$. Consider f, g as identical if $f(x) \equiv g(x)$, except for an x-set of μ-measure 0. Define*

$$(f, g)_S = (f, g)_{S,\Gamma,\mu} = \int_S f(x)\, \overline{g(x)}\, d_\mu x.$$

LEMMA 3.2.2. *The space $\mathfrak{H}_S$ fulfills the postulates A, B, C, E of* (**4**), *pp. 64–66. (This includes the existence and finiteness of the integral which defines $(f, g)_S$ for $f, g \in \mathfrak{H}_S$.) Thus $\mathfrak{H}_S$ is a finite-dimensional Euclidean or a Hilbert space.*

PROOF: Existence and finiteness of $\int_S f(x)\, \overline{g(x)}\, d_\mu x$: Literally as in (**4**), p. 109.

Ad A, B, E: Literally as in (**4**), pp. 109, 111.

Ad C: Same as in (**4**), pp. 109–111, except that the system of neighborhoods in which occurs there (Ω is a metric space which plays the rôle of our S), is to be replaced by the following sets in S (all $\in \Gamma$): $T^{(i)} T_j$, $i, j = 1, 2, \cdots$, where the $T^{(1)}$, $T^{(2)}$, $\cdots$ are the sets from (iii) in Definition 3.2.4, and the T_1, T_2, $\cdots$ are all sets which obtain by any (finite) number of applications of the operations $M - N$, $M + N$, $M \cdot N$ to the elements of a finite or enumerably infinite basis Δ of Γ in the sense of Definitions 3.2.2, 3.2.3.—

It may be well to recall at this stage the examples a) and b) of (**1**), p. 193. (The sets $T^{(1)}$, $T^{(2)}$, $\cdots$ given there serve both for (iii) in Definition 3.2.4, and to construct Δ, Γ by $\Delta = (T^{(1)}, T^{(2)}, \cdots)$, $\Gamma = \Gamma(\Delta)$. Γ turns out to be in the example a) the system of all sub-sets of S and in the example b) the system of all relative Borel sets in S in the usual topological sense.) The example a), with a finite set $S = (x_1, \cdots, x_n)$ shows in particular that $\mathfrak{H}_S$ may actually be a finite-dimensional Euclidean space.

We conclude this section by proving the "differentiation-theorem" of Lebesgue-Nikodym (Theorem VII below). It would be possible to carry out

most of what we want also without this theorem, but such a procedure would necessitate qualifications and restrictions in the hypotheses of the lemmas and theorems which follow.

LEMMA 3.2.3. *Consider two* Γ*-measures* μ, ν *in* S. *Then we have*:

(i) *If* $\mu(M) = 0$ *implies* $\nu(M) = 0$ *then there exists a* Γ*-measurable function* $\kappa(x) \geqq 0$, *such that for all* Γ*-measurable (complex) functions* $f(x)$ *and all sets* $M \in \Gamma$

$$\int_M f(x)\, d_\mu x = \int_M f(x)\kappa(x)\, d_\nu x\,.$$

(*This is meant to include that if either side is defined and finite, then both are.*)
(ii) *The* $\kappa(x)$ *of* (i) *is unique, up to* x*-sets of* μ*-measure* 0. *This is even the case if* (i) *is restricted to* $f(x) \equiv 1$, $M \in \Gamma$.
(iii) *The equivalence of* $\mu(M) = 0$ *to* $\nu(M) = 0$ *is equivalent to having always* $\kappa(x) > 0$ *in* (i), *except for an* x*-set of* μ*-measure* 0.

PROOF: Ad (i): Consider the $T^{(1)}, T^{(2)}, \cdots$ of (iii) in Definition 3.2.4 first for μ: $T_1^{(1)}, T_1^{(2)}, \cdots$ and then for ν: $T_2^{(1)}, T_2^{(2)}, \cdots$. Let $T_0^{(1)}, T_0^{(2)}, \cdots$ be the $T_1^{(i)} \cdot T_2^{(j)}$, $i, j = 1, 2, \cdots$, arranged as a sequence. They will then do for the Γ-measure defined by

$$\rho(M) = \mu(M) + \nu(M) \qquad (M \in \Gamma) \tag{48}$$

in the sense of (iii) in Definition 3.2.4. The same is true for $T_{00}^{(1)}, T_{00}^{(2)}, \cdots$, if we define $T_{00}^{(i)} = T_0^{(i)} - (T_0^{(1)} + \cdots + T_0^{(i-1)})$, $i = 1, 2, \cdots$. And these have the further property that $i \neq j$ implies $T_{00}^{(i)} \cdot T_{00}^{(j)} = 0$.

Let us form

$$L_i(f) = \int_{T_{00}^{(i)}} \overline{f(x)}\, d_\nu x \qquad \text{for any } f \in \mathfrak{H}_{S,\Gamma,\rho}. \tag{49}$$

(Observe: ν in the integral, but ρ in $\mathfrak{H}_{S,\Gamma,\rho}$!) Since

$$\int_{T_{00}^{(i)}} |f(x)|^2\, d_\nu x \leqq \int_{T_{00}^{(i)}} |f(x)|^2\, d_\rho x = \int_S |f(x)|^2\, d_\rho x = \|f\|^2_{S,\Gamma,\rho}$$

and

$$\int_{T_{00}^{(i)}} 1^2\, d_\nu x = \nu(T_{00}^{(i)})$$

are both finite, the integral in (49) is (by Schwarz's well-known inequality) existing and finite, and even

$$|L_i(f)| = \left|\int_{T_{00}^{(i)}} f(x)\, d_\nu x\right| \leqq \left[\int_{T_{00}^{(i)}} |f(x)|^2\, d_\nu x \cdot \int_{T_{00}^{(i)}} 1^2\, d_\nu x\right]^{\frac{1}{2}}$$

$$\leqq [\nu(T_{00}^{(i)})]^{\frac{1}{2}} \cdot \|f\|_{S,\Gamma,\rho},$$

$$|L_i(f)| \leqq C_i \cdot \|f\|_{S,\Gamma,\rho} \qquad (C_i = [\nu(T_{00}^{(i)})]^{\frac{1}{2}}). \tag{50}$$

$L_i(f)$ is by (49) a conjugate linear functional in $\mathfrak{H}_{S,\Gamma,\rho}$, and by (50) a continuous one. Hence a lemma of F. Riesz applies to it (cf. e.g. (4), p. 94, footnote 52): There exists a $\kappa_i' \in \mathfrak{H}_{S,\Gamma,\rho}$, such that $L_i(f) = (\kappa_i', f)$ that is

$$\int_{T_{00}^{(i)}} \overline{f(x)}\, d_\nu x = \int_S \overline{f(x)}\, \kappa_i'(x)\, d_\rho x = \int_S \overline{f(x)}\, \kappa_i'(x)\, d_\mu x + \int_S \overline{f(x)}\, \kappa_i'(x)\, d_\nu x .$$

Consider now a Γ-measurable $f(x) \geqq 0$, which is $= 0$ for $x \notin T_{00}^{(i)}$, and bounded in S. Then $f \in \mathfrak{H}_{S,\Gamma,\rho}$ (because $\rho(T_{00}^{(i)})$ is finite!), and the above equation may be written

$$\int_{T_{00}^{(i)}} f(x)(1 - \kappa_i'(x))\, d_\nu x = \int_{T_{00}^{(i)}} f(x)\, \kappa_i'(x)\, d_\mu x . \tag{51}$$

Since it does not matter in (51), what $f(x)$ does for $x \notin T_{00}^{(i)}$ we can drop the requirement of $f(x) = 0$ for $x \notin T_{00}^{(i)}$.

Taking the real part of both sides of (51) replaces (since $f(x)$ is real) $\kappa_i'(x)$ by $\Re\kappa_i'(x)$. Hence we can assume that $\kappa_i'(x)$ is real. Now put

$$f(x) \equiv \theta_M(x) \begin{cases} = 1 \text{ for } x \in M \\ = 0 \text{ otherwise} \end{cases}$$

first with $M = M_i' = T_{00}^{(i)} \cdot (x;\ \kappa_i'(x) < 0)$, and then with $M = M_i'' = T_{00}^{(i)} \cdot (x;\ \kappa_i'(x) \geqq 1)$. In both cases $f(x)$ meets the requirements of (51). The left side of (51) becomes $\geqq 0$ and $\leqq 0$ respectively, and so we obtain

$$\int_{M_i'} \kappa_i'(x)\, d_\mu x \geqq 0, \qquad \int_{M_i''} \kappa_i'(x)\, d_\mu x \leqq 0 .$$

But $\kappa'(x) < 0$ for all $x \in M_i'$ and $\kappa'(x) > 0$ for all $x \in M_i''$; hence necessarily $\mu(M_i') = \mu(M_i'') = 0$. This also implies $\nu(M_i') = \nu(M_i'') = 0$. Hence we can change $\kappa_i'(x)$ arbitrarily in $M_i' + M_i''$ without affecting the validity of (51). We redefine $\kappa_i'(x) = 0$ for $x \in M_i' + M_i''$. So we have

$$0 \leqq \kappa_i'(x) < 1 \qquad (x \in T_{00}^{(i)}). \tag{52}$$

So the integrands on both sides of (51) are $\geqq 0$. Therefore if $f(x)$ is merely Γ-measurable and $\geqq 0$, we can put

$$f_n(x) \left\{ \begin{array}{l} = f(x) \text{ for } f(x) \leqq n \\ = 0 \text{ otherwise} \end{array} \right| ,$$

replace $f(x)$ by $f_n(x)$ in (51), and apply $\lim_{n\to\infty}$: Then (51) for the original $f(x)$ results. In other words: (51) holds for all Γ-measurable $f(x) \geqq 0$. (Of course then both sides of (51) may be infinite.)

Now put $\kappa_i(x) \equiv \dfrac{\kappa_i'(x)}{1 - \kappa_i'(x)}$. Then (52) gives

$$\kappa_i(x) \geqq 0 \qquad (x \in T_{00}^{(i)}). \tag{53}$$

We can replace $f(x)$ in (51) by $\dfrac{f(x)}{1-\kappa_i'(x)}$; then (51) becomes:

$$\int_{T_{00}^{(i)}} f(x)\, d_\nu x = \int_{T_{00}^{(i)}} f(x)\, \kappa_i(x)\, d_\mu x\,, \tag{54}$$

again for all Γ-measurable $f(x) \geqq 0$. Now define $\kappa(x) \equiv \kappa_i(x)$ for $x \in T_{00}^{(i)}$, $i = 1, 2, \cdots$. Then (53) gives

$$\kappa(x) \geqq 0 \qquad (x \in S). \tag{55}$$

And adding (54) over all $i = 1, 2, \cdots$ gives

$$\int_S f(x)\, d_\nu x = \int_S f(x)\, \kappa(x)\, d_\mu x\,. \tag{56}$$

Consider now any Γ-measurable (complex) $f(x)$. Apply (56) to $|f(x)|$ (which is $\geqq 0$): It shows that if either of $\int_S |f(x)|\, d_\nu x$, $\int_S |f(x)|\, \kappa(x) d_\mu x$ is finite, then both are. That is ($f(x)$ is complex!): If either of $\int_S f(x) d_\nu x$, $\int_S f(x)\kappa(x) d_\mu x$ is defined and finite, then both are. Combining this with the validity of (56) for Γ-measurable $f(x) \geqq 0$ shows: (56) holds for all Γ-measurable (complex) $f(x)$, this being meant to include, that if either side of (56) is defined and finite, then both are.

Now define $f'(x)\begin{cases} = f(x) \text{ for } x \in M \\ = 0 \text{ otherwise} \end{cases}$, $M \in \Gamma$. Then (56) becomes (in the same sense as above)

$$\int_M f(x)\, d_\nu x = \int_M f(x)\, \kappa(x)\, d_\mu x\,. \tag{57}$$

And (55), (57) complete the proof of (i).

Ad (ii): Assume that $\kappa'(x)$, $\kappa''(x)$ both meet the requirements of (i), with the restriction made in (ii). Consider $T^{(1)}, T^{(2)}, \cdots$ of (iii) in Definition 3.2.4 for ν, and put $M'_{i,n} = T^{(i)} \cdot (x;\ \kappa'(x) < \kappa''(x) \leqq n)$. Consider the function $f(x) \equiv 1$ and $M = M'_{i,n}$. Then we have $\nu(M'_{i,n}) = \int_{M'_{i,n}} \kappa'(x)\, d_\mu x = \int_{M'_{i,n}} \kappa''(x)\, d_\mu x$. Since $\nu(M'_{i,n}) \leqq \nu(T^{(i)})$ is finite, we can infer $\int_{M'_{i,n}} (\kappa''(x) - \kappa'(x)) d_\mu x = 0$. But $\kappa''(x) - \kappa'(x) > 0$ for $x \in M'_{i,n}$, hence we have $\mu(M'_{i,n}) = 0$. Put $M' = \sum_{i,n=1}^{\infty} M'_{i,n} = (x;\ \kappa'(x) < \kappa''(x))$, then this gives $\mu(M') = 0$. By symmetry in $\kappa'(x)$, $\kappa''(x)$ we have similarly for $M'' = (x;\ \kappa'(x) > \kappa''(x))$ $\mu(M'') = 0$. And so for $M = M' + M'' = (x;\ \kappa'(x) \neq \kappa''(x))$ again $\mu(M) = 0$. That is: $\kappa'(x) \equiv \kappa''(x)$, except for an x-set of μ-measure 0 (namely: M).

Ad (iii): *Sufficiency*: Assume $\kappa(x) > 0$, except for an x-set of μ-measure 0. $\nu(M) = 0$ implies $\int_M \kappa(x) d_\mu x = 0$, and since $\kappa(x) > 0$ for $x \in M$ (except for an

x-set of μ-measure 0), so $\mu(M) = 0$. Hence $\mu(M) = 0$ and $\nu(M) = 0$ are equivalent.

Necessity: Assume that $\mu(M) = 0$ is equivalent to $\nu(M) = 0$. Put

$$f(x) \equiv \theta_M(x) \begin{cases} = 1 \text{ for } x \in M \\ = 0 \text{ otherwise} \end{cases},$$

$M = (x; \kappa(x) = 0)$. Then we obtain $\nu(M) = \int_M \kappa(x) d_\mu x = 0$, hence $\mu(M) = 0$ (by the above assumption). So we have $\kappa(x) > 0$, except for an x-set of μ-measure 0 (namely: M).

DEFINITION 3.2.7. *We denote the unique (up to x-sets of μ-measure 0!) $\kappa(x)$ of* (i), (ii) *in Lemma* 3.2.3 *by* $\frac{d_\nu}{d_\mu}(x)$.

THEOREM VII. *Let μ be a fixed Γ-measure in S. Then the relation*

$$\kappa(x) \equiv \frac{d_\nu}{d_\mu}(x)$$

establishes a one-to-one correspondence between all Γ-measurable functions $\kappa(x) \geqq 0$ (up to x-sets of μ-measure 0!) on one hand, and all Γ-measures ν, for which $\mu(M) = 0$ implies $\nu(M) = 0$, on the other. The inverse relation is

$$\nu(M) = \int_M \kappa(x)\, d_\mu x.$$

PROOF: Given $\nu(M)$, as described above, the existence and uniqueness of this $\kappa(x)$ was already stated in (i), (ii) in Lemma 3.2.3 and in Definition 3.2.7. So we must only consider the case where $\kappa(x)$, as described above, is given.

Consider the $T^{(1)}, T^{(2)}, \cdots$ of (iii) in Definition 3.2.4, and let $T_0^{(1)}, T_0^{(2)}, \cdots$ be the $T^{(i)} \cdot (x; \kappa(x) \leqq n)$, $i, n = 1, 2, \cdots$, arranged as a sequence. They will then do for the Γ-measure defined by

$$\nu(M) = \int_M \kappa(x)\, d_\mu x \qquad (M \in \Gamma) \tag{58}$$

in the sense of (iii) in Definition 3.2.4. It is also clear by (58) that $\mu(M) = 0$ implies $\nu(M) = 0$. Hence we can form $\frac{d_\nu}{d_\mu}(x)$.

Put $f(x) \equiv 1$, then we have $\nu(M) = \int_M d_\nu x = \int_M \frac{d_\nu}{d_\mu}(x)\, d_\mu x$. So (58) gives

$$\int_M \frac{d_\nu}{d_\mu}(x)\, d_\mu x = \int_M \kappa(x)\, d_\mu x \qquad (M \in \Gamma).$$

Hence by (ii) in Lemma 3.2.3 $\frac{d_\nu}{d_\mu}(x) \equiv \kappa(x)$, except for an x-set of μ-measure 0. This, together with (58), completes the proof.

COROLLARY. (i) $\frac{d_\mu}{d_\mu}(x) \equiv 1$.

(ii) $\frac{d_\nu}{d_\mu}(x)\frac{d_\rho}{d_\nu}(x) \equiv \frac{d_\rho}{d_\mu}(x)$ *if* $\mu(M) = 0$ *implies* $\nu(M) = 0$, *and* $\nu(M) = 0$ *implies* $\rho(M) = 0$.

PROOF: (i), (ii) are both immediate on the basis of the characterization of the $\frac{d_\nu}{d_\mu}(x)$ in (i), (ii) in Lemma 3.2.3.

3.3. Let S, Γ, μ be as in the preceding section. We now describe a certain relationship between S and a group $\mathfrak{G}$.

DEFINITION 3.3.1. $\mathfrak{G}$ *is an* S*-group if it has the following properties*:
(i) $\mathfrak{G}$ *is a "group," that is, a "composition"* $\alpha\beta$, *an "inverse"* α^{-1}, *and a "unit"* $\mathbf{1}$ *are defined in* $\mathfrak{G}$ *with these properties*:

$$(\alpha\beta)\gamma = \alpha(\beta\gamma), \qquad \alpha\alpha^{-1} = \alpha^{-1}\alpha = \mathbf{1}, \qquad \alpha\mathbf{1} = \mathbf{1}\alpha = \alpha.$$

(*So* $\mathfrak{G}$ *is "associative," but not necessarily "commutative."*)
(ii) *Every* $\alpha \in \mathfrak{G}$ *defines a certain one-to-one mapping of* S *on itself*:

$$x \rightleftarrows x\alpha \qquad (x, x\alpha \in S).$$

These mappings "represent" $\mathfrak{G}$, *that is*

$$(x\alpha)\beta = x(\alpha\beta) \qquad (x \in S, \alpha, \beta \in \mathfrak{G}).$$

(iii) $\mathfrak{G}$ *is finite or enumerably infinite.*—
One sees immediately (by (ii) above) that $x\mathbf{1} \equiv x$ and that $x \rightleftarrows x\alpha^{-1}$ is the inverse mapping to $x \rightleftarrows x\alpha$.

We shall, as usual, denote by $M\alpha$ the set $(x\alpha; x \in M)$ $(M, M\alpha \subset S)$.

DEFINITION 3.3.2. *An* S*-group* $\mathfrak{G}$ *is an "*S, Γ*-group" if for every* $\alpha \in \mathfrak{G}$ $M \in \Gamma$ *implies* $M\alpha \in \Gamma$. *Hence if* μ *is a* Γ*-measure in* S, *then the* μ_α *defined by*

$$\mu_\alpha(M) \equiv \mu(M\alpha)$$

is also a Γ*-measure in* S.

DEFINITION 3.3.3. *An* S, Γ*-group* $\mathfrak{G}$ *is an "*S, Γ, μ*-group"* (μ *a* Γ*-measure in* S), *if for every* $\alpha \in \mathfrak{G}$ $\mu(M) = 0$ *implies* $\mu(M\alpha) = 0$.

LEMMA 3.3.1. *Let* $\mathfrak{G}$ *be an* S, Γ, μ*-group. Then we have*:
(i) $\mu(M) = 0$ *and* $\mu_\alpha(M) = 0$ (*that is*: $\mu(M\alpha) = 0$) *are equivalent. Hence we can form* $\frac{d_{\mu_\alpha}}{d_\mu}(x)$.
(ii) *Each one of the following formulae holds, except for* x*-sets of* μ*-measure* 0:

$$\frac{d_{\mu_\alpha}}{d_\mu}(x) > 0,$$

$$\frac{d_{\mu_1}}{d_\mu}(x) \equiv \frac{d_\mu}{d_\mu}(x) \equiv 1,$$

$$\frac{d_{\mu_\beta}}{d_\mu}(x\alpha)\frac{d_{\mu_\alpha}}{d_\mu}(x) \equiv \frac{d_{\mu_{\alpha\beta}}}{d_\mu}(x).$$

(iii) $\frac{d_{\mu_\alpha}}{d_\mu}(x)$ *is uniquely characterized, up to an* x*-set of* μ*-measure* 0, *by this property*

$$\int_{M\alpha^{-1}} f(x\alpha)\,\frac{d_{\mu_\alpha}}{d_\mu}(x)\,d_\mu x = \int_M f(x)\,d_\mu x.$$

($f(x)$ *is any* Γ*-measurable function,* $M \in \Gamma$, *everything in the same sense as in* (i) *in Lemma* 3.2.3).

PROOF: We prove these statements in a somewhat changed order.

Ad (i): $\mu(M) = 0$ implies $\mu(M\alpha) = 0$, and $\mu(M\alpha) = 0$ implies $\mu(M\alpha\cdot\alpha^{-1}) = 0$, $\mu(M) = 0$. Hence $\mu(M) = 0$ is equivalent to $\mu(M\alpha) = 0$, that is to $\mu_\alpha(M) = 0$.

Ad (ii), *first formula*: Immediate by (i) above and (iii) in Lemma 3.2.3.

Ad (iii): According to Definition 3.2.7, $\frac{d_{\mu_\alpha}}{d_\mu}(x)$ is characterized, up to an x-set of μ-measure 0, by

$$\int_M f(x)\,d_{\mu_\alpha}x = \int_M f(x)\,\frac{d_{\mu_\alpha}}{d_\mu}(x)\,d_\mu x \tag{59}$$

($f(x)$ Γ-measurable, $M \in \Gamma$). Now obviously $\int_M f(x)d_{\mu_\alpha}x = \int_{M\alpha} f(x\alpha^{-1})d_\mu x$. So (59) becomes

$$\int_{M\alpha} f(x\alpha^{-1})\,d_\mu x = \int_M f(x)\,\frac{d_{\mu_\alpha}}{d_\mu}(x)\,d_\mu x. \tag{60}$$

Replacement of $M, f(x)$ by $M\alpha^{-1}, f(x\alpha)$ in (60) gives the desired formula.

Ad (ii), *second formula*: Coincides with (i) in the Corollary after Theorem VII.

Ad (iii), *third formula*: Two applications of (iii) above show, that $\frac{d_{\mu_\beta}}{d_\mu}(x\alpha)\frac{d_{\mu_\alpha}}{d_\mu}(x)$ satisfies the characterization of $\frac{d_{\mu_{\alpha\beta}}}{d_\mu}(x)$ by (iii) above.

DEFINITION 3.3.4. *An* S, Γ, μ*-group* $\mathfrak{G}$ *is "free" if it has this property*:

If $\alpha \in \mathfrak{G}$ *is* $\neq \mathbf{1}$ *then* $x\alpha = x$ *holds only for an* x*-set of* μ*-measure* 0.

DEFINITION 3.3.5. *An* S, Γ, μ*-group* $\mathfrak{G}$ *is "ergodic" if it has this property*:

If $M \in \Gamma$ *is such that for every* $\alpha \in \mathfrak{G}$ *the difference of the sets* M, $M\alpha$ (*i.e. the set* $(M + M\alpha) - (M \cdot M\alpha)$) *has* μ*-measure* 0, *then either* $\mu(M) = 0$ *or* $\mu(\mathfrak{G} - M) = 0$.—

Cf. the comment in (**1**), p. 195, bottom of the page, concerning this notion of ergodicity.

LEMMA 3.3.2. *An* S, Γ, μ*-group* $\mathfrak{G}$ *is ergodic if and only if it has the following property*:

If $f(x)$ *is a* Γ*-measurable function such that for every* $\alpha \in \mathfrak{G}$ $f(x\alpha) \equiv f(x)$ (*except for* x*-sets of* μ*-measure* 0), *then* $f(x) \equiv a$ (*constant!, except for an* x*-set of* μ*-measure* 0).

PROOF: *Sufficiency*: If $f(x)$ runs only over characteristic functions

$$f(x) \equiv \theta_M(x)\left\{\begin{array}{l} = 1 \text{ for } x \in M \\ = 0 \text{ otherwise}\end{array}\right|,$$

then this coincides with Definition 3.3.5.

Necessity: Since we may consider $\Re f(x)$, $\Im f(x)$ instead of $f(x)$ we may assume that $f(x)$ is real. And since we may consider $\frac{f(x)}{1+|f(x)|}$ instead of $f(x)$, we may even assume that

$$-1 < f(x) < 1. \tag{61}$$

For any real a the set $M_a = (x; f(x) < a)$ has clearly the property used in Definition 3.3.5; hence either $\mu(M_a) = 0$ or $\mu(\mathfrak{G} - M_a) = 0$. Let Σ be the set of all a with $\mu(M_a) = 0$. By (61) $-1 \in \Sigma$, $1 \notin \Sigma$, and as for $a \leqq b$ $M_a \subset M_b$, so in this case $b \in \Sigma$ implies $a \in \Sigma$. Hence Σ possesses a least upper bound, say a_0 and $a_0 - \frac{1}{n} \in \Sigma$, $a_0 + \frac{1}{n} \notin \Sigma$ for $n = 1, 2, \cdots$.

Hence the sets $(x; f(x) < a_0) = \sum_{n=1}^{\infty} M_{a_0-1/n}$ and $(x; f(x) > a_0) = \sum_{n=1}^{\infty} (\mathfrak{G} - M_{a_0+1/n})$ have μ-measure 0, and so has their sum, the set $(x; f(x) \neq a_0)$. That is: $f(x) \equiv a_0$, except for an x-set of μ-measure 0.—

Our present notion of a free, ergodic, S, Γ, μ-group corresponds to what was called an ergodic m-group (for S, μ) in (**1**), page 195, Definition 12.1.5, but there the stronger requirement $\mu(M) = \mu(M\alpha)$ replaced our present requirement that $\mu(M) = 0$ imply $\mu(M\alpha) = 0$.

This situation motivates the following definition:

DEFINITION 3.3.6. *An S, Γ, μ-group $\mathfrak{G}$ is "measurable" if there exists a Γ-measure ν with the following properties*:

(i) *$\mu(M) = 0$ is equivalent to $\nu(M) = 0$.*

(ii) *Always $\nu(M) = \nu(M\alpha)$.*—

We conclude this section by proving two properties of the above notions.

LEMMA 3.3.3. *The properties of the Γ-measure ν given in Definition* 3.3.6, *are equivalent to the following properties of its $\kappa(x) \equiv \frac{d_\nu}{d_\mu}(x)$ (remember Theorem* VII!), *always except for x-sets of μ-measure* 0:

$$\kappa(x) > 0. \tag{i}$$

$$\frac{\kappa(x)}{\kappa(x\alpha)} \equiv \frac{d_{\mu_\alpha}}{d_\mu}(x). \tag{ii}$$

PROOF: (i) is equivalent to (i) in Definition 3.3.6: Coincides with (iii) in Lemma 3.2.3.

(ii) is equivalent to (ii) in Definition 3.3.6: $\frac{d_\nu}{d_\mu}(x) \equiv \kappa(x)$, hence clearly $\frac{d_{\nu_\alpha}}{d_{\mu_\alpha}}(x) = \kappa(x\alpha)$. So (ii) in the Corollary to Theorem VII gives (by evaluating $\frac{d_{\nu_\alpha}}{d_\mu}(x)$ in two ways):

$$\kappa(x)\frac{d_{\nu_\alpha}}{d_\nu}(x) \equiv \kappa(x\alpha)\frac{d_{\mu_\alpha}}{d_\mu}(x). \tag{62}$$

Now (ii) in Definition 3.3.6 means $\nu = \nu_\alpha$, that is, by (i) in the Corollary to Theorem VII $\frac{d\nu_\alpha}{d\nu}(x) \equiv 1$. Hence (62) becomes the desired equation $\frac{\kappa(x)}{\kappa(x\alpha)} \equiv \frac{d\mu_\alpha}{d\mu}(x)$.

LEMMA 3.3.4. *Let $\mathfrak{G}$ be an ergodic S, Γ, μ-group. Then the Γ-measure ν of Definition 3.3.6 is uniquely determined, up to a constant factor $a > 0$.*

PROOF: Let ν', ν'' be two such Γ-measures. Form their $\kappa'(x) = \frac{d\nu'}{d\mu}(x)$, $\kappa''(x) = \frac{d\nu''}{d\mu}(x)$, in accordance with Lemma 3.3.3. Then (ii) in Lemma 3.3.3 gives $\kappa'(x)/\kappa'(x\alpha^{-1}) \equiv \kappa''(x)/\kappa''(x\alpha^{-1})$, $\kappa'(x)/\kappa''(x) \equiv \kappa'(x\alpha^{-1})/\kappa''(x\alpha^{-1})$. Put $f(x) \equiv \kappa'(x)/\kappa''(x)$ and replace α by α^{-1}, then we have: $f(x) \equiv f(x\alpha)$. All this, of course, holds with the exception of x-sets of μ-measure 0. Now Lemma 3.3.2 gives $f(x) \equiv a$, $\kappa'(x) \equiv a\kappa''(x)$(in the same sense) with a constant a. Hence $\nu'(M) = a\nu''(M)$, and clearly $a > 0$.

3.4. DEFINITION 3.4.1. (i) *Let $\mathfrak{H}^{\mathfrak{G}}$ be the set of all (complex-valued) functions $f(\alpha)$ ($\alpha \in \mathfrak{G}$) with a finite $\sum_{\alpha \in \mathfrak{G}} |f(\alpha)|^2$. (Remember that $\mathfrak{G}$ is finite or enumerably infinite!) Define*

$$(f, g)^{\mathfrak{G}} = \sum_{\alpha \in \mathfrak{G}} f(\alpha)\overline{g(\alpha)}.$$

(ii) *Let $\mathfrak{H}_S^{\mathfrak{G}} = \mathfrak{H}_{S,\Gamma,\mu}^{\mathfrak{G}}$ be the set of all (complex-valued) functions $\mathcal{F}(x, \alpha)$ ($x \in S$, $\alpha \in \mathfrak{G}$) which are Γ-measurable in x for every $\alpha \in \mathfrak{G}$ and with a finite*

$$\sum_{\alpha \in \mathfrak{G}} \int_S |\mathcal{F}(x, \alpha)|^2 \, d_\mu x.$$

Consider $\mathcal{F}$, $\mathcal{H}$ as identical if $\mathcal{F}(x, \alpha) \equiv \mathcal{H}(x, \alpha)$ except for an x-set of μ-measure 0 for every $\alpha \in \mathfrak{G}$. Define

$$(\mathcal{F}, \mathcal{H})_S^{\mathfrak{G}} = (\mathcal{F}, \mathcal{H})_{S,\Gamma,\mu}^{\mathfrak{G}} = \sum_{\alpha \in \mathfrak{G}} \int_S \mathcal{F}(x, \alpha)\overline{\mathcal{H}(x, \alpha)} \, d_\mu x.$$

LEMMA 3.4.1. (i) *The spaces $\mathfrak{H}^{\mathfrak{G}}$, $\mathfrak{H}_S^{\mathfrak{G}}$ fulfill the postulates* A, B, C, E *of* (4), *pp.* 64–66. *(This includes the existence and finiteness of the sums and integrals which define $(f, g)^{\mathfrak{G}}$, $(\mathcal{F}, \mathcal{H})_S^{\mathfrak{G}}$.) Thus $\mathfrak{H}^{\mathfrak{G}}$, $\mathfrak{H}_S^{\mathfrak{G}}$ are finite dimensional Euclidean or Hilbert spaces.*

(ii) *Using the complete, normalized, orthogonal set of all functions*

$$\varphi_{\alpha_0}(\alpha)\begin{cases} = 1 & \text{for } \alpha = \alpha_0 \\ = 0 & \text{otherwise} \end{cases}, \qquad \alpha_0 \in \mathfrak{G},$$

in $\mathfrak{H}^{\mathfrak{G}}$ *and setting up the correspondence*

$$\mathcal{F}(x, \alpha) \sim < \mathcal{F}(x, \bar{\alpha}_1), \quad \mathcal{F}(x, \bar{\alpha}_2), \cdots >$$

($\bar{\alpha}_1, \bar{\alpha}_2, \cdots$ *an enumeration of* $\mathfrak{G}$, *so that the above* α_0 *runs over* $\bar{\alpha}_1, \bar{\alpha}_2, \cdots$) $\mathfrak{H}_s^{\mathfrak{G}}$ *is isomorphic with* $\mathfrak{H}^{\mathfrak{G}} \otimes \mathfrak{H}_s$ *in the sense of* (1), *p.* 135, *Definition* 2.4.1 *and Lemma* 2.4.1.

In what follows we use the more symmetric notation

$$\mathcal{F}(x, \alpha) \sim < \mathcal{F}(x, \alpha); \quad \alpha \in \mathfrak{G} >.$$

PROOF: We prove these statements in a somewhat changed order.

Ad (i): Character of $\mathfrak{H}^{\mathfrak{G}}$: Put $f(\bar{\alpha}_n) = \xi_n$, $n = 1, 2, \cdots$ ($\bar{\alpha}_1, \bar{\alpha}_2, \cdots$ an enumeration of $\mathfrak{G}$!), and $f(\alpha) \sim \{\xi_1, \xi_2, \cdots\}$ then our definitions coincide with the original definition of a finite dimensional Euclidean or a Hilbert space ((1), p. 69, lower part of the page).

Ad (ii): Immediate, by comparing our Definition 3.4.1 with (1), p. 135, Definition 2.4.1 and Lemma 2.4.1.

Ad (i): Character of $\mathfrak{H}_s^{\mathfrak{G}}$: Considering what we showed about $\mathfrak{H}^{\mathfrak{G}}$ above, and about $\mathfrak{H}_s$ in Lemma 3.2.2, this follows from (ii) above, since $\mathfrak{H}_s^{\mathfrak{G}}$ is isomorphic to $\mathfrak{H}^{\mathfrak{G}} \otimes \mathfrak{H}_s$.—

We are now in a position to begin a systematic investigation of certain operators in $\mathfrak{H}_s$ and $\mathfrak{H}_s^{\mathfrak{G}}$.

3.5. Throughout what follows, $\mathfrak{G}$ will be a fixed S, Γ, μ-group. We assume from now on throughout §§3.5–3.7, that $\mathfrak{G}$ is free and ergodic, but not that it is measurable.

We begin by constructions in $\mathfrak{H}_s$.

LEMMA 3.5.1. *Define the operators*:

(i) $U_\alpha f(x) \equiv \left[\frac{d_{\mu_\alpha}}{d_\mu}(x)\right]^{\frac{1}{2}} f(x\alpha)$ *for any* $\alpha \in \mathfrak{G}$.

(ii) $L_{\varphi(x)} f(x) \equiv \varphi(x) f(x)$ *for any* (*complex-valued*) Γ-*measurable, bounded function* $\varphi(x)$.

These operators are bounded, U_α *is even unitary.*

PROOF: U_α is bounded and unitary: By (iii) in Lemma 3.3.1

$$\int_S |U_\alpha f(x)|^2 d_\mu x = \int_S |f(x\alpha)|^2 \frac{d_{\mu_\alpha}}{d_\mu}(x) d_\mu x = \int_S |f(x)|^2 d_\mu x.$$

So $f \in \mathfrak{H}_s$ implies $U_\alpha f \in \mathfrak{H}_s$ and $\| U_\alpha f \|_s = \| f \|_s$.

$L_{\varphi(x)}$ is bounded: We have $|\varphi(x)| \leqq C$ for all $x \in S$ and a suitable $C > 0$. Hence

$$\int_S |L_{\varphi(x)} f(x)|^2 d_\mu x = \int_S |\varphi(x)|^2 |f(x)|^2 d_\mu x \leqq C^2 \int_S |f(x)|^2 d_\mu x.$$

So $f \in \mathfrak{H}_s$ implies $L_{\varphi(x)} f \in \mathfrak{H}_s$, and $\| L_{\varphi(x)} f \|_s \leqq C \| f \|_s$.

LEMMA 3.5.2. (i) $U_1 = 1$.
(ii) $U_{\alpha^{-1}} = U_\alpha^{-1} = U_\alpha^*$.
(iii) $U_{\alpha\beta} = U_\alpha U_\beta$.

PROOF: We prove these statements in a somewhat changed order:
Ad (i), (iii): Immediate by (ii) in Lemma 3.3.1.
Ad (ii): (i), (iii) above give $U_{\alpha^{-1}} = U_\alpha^{-1}$. And since U_α is unitary, so $U_\alpha^{-1} = U_\alpha^*$

LEMMA 3.5.3. *Let $\mathfrak{L}$ be the set of all operators $L_{\varphi(x)}$ from* (ii) *in Lemma* 3.5.1. *Then $\mathfrak{L}$ is a ring, and $\mathfrak{L} = \mathfrak{L}'$.*

PROOF: As $\mathfrak{L}'$ is necessarily a ring it suffices to prove $\mathfrak{L} = \mathfrak{L}'$. As $L_{\varphi(x)}L_{\psi(x)} = L_{\varphi(x)\psi(x)}$, so $L_{\varphi(x)}$ commutes with every $L_{\psi(x)}$. As $L_{\psi(x)}^* = L_{\overline{\psi(x)}}$, so $L_{\varphi(x)}$ also commutes with every $L_{\psi(x)}^*$. Thus $L_{\varphi(x)} \in \mathfrak{L}'$, that is, $\mathfrak{L} \subset \mathfrak{L}'$. So we must only prove $\mathfrak{L}' \subset \mathfrak{L}$.

Consider therefore an $A \in \mathfrak{L}'$. Then A commutes with every $L_{\varphi(x)}$, that is $AL_{\varphi(x)}f = L_{\varphi(x)}Af$. That is

$$A(\varphi(x)f(x)) \equiv \varphi(x)Af(x). \tag{63}$$

The assumptions are: $\varphi(x)$, $f(x)$ Γ-measurable, $\varphi(x)$ bounded, $\int_S |f(x)|^2 d_\mu x$ finite. (63), as well as all subsequent relations in this proof, holds with the exception of an x-set of μ-measure 0.

Consider the $T^{(1)}, T^{(2)}, \cdots$ of (iii) in Definition 3.2.4; we can replace them by $T_0^{(1)}, T_0^{(2)}, \cdots$ where $T_0^{(i)} = T^{(i)} - (T^{(1)} + \cdots + T^{(i-1)})$, and then we have also that $i \neq j$ implies $T_0^{(i)}, T_0^{(j)} = \Theta$. Define

$$e_i(x) \equiv \theta_{T_0^{(i)}}(x) \begin{cases} = 1 & \text{for } x \in T_0^{(i)} \\ = 0 & \text{otherwise} \end{cases}.$$

Then $\int_S |e_i(x)|^2 d_\mu x = \mu(T_0^{(i)})$ is finite, and so we can put $f(x) \equiv e_i(x)$ in (63). We obtain:

$$A(\varphi(x)e_i(x)) \equiv (Ae_i(x))\varphi(x). \tag{64}$$

As $e_i(x)$ is bounded, we can also put $\varphi(x) \equiv e_i(x)$ in (64), and so we obtain $A(e_i(x)) \equiv (Ae_i(x))e_i(x)$ that is:

$$Ae_i(x) \equiv 0 \quad \text{for} \quad x \notin T_0^{(i)} \tag{65}$$

Now define $\rho(x) = Ae_i(x)$ for $x \in T_0^{(i)}$, $i = 1, 2, \cdots$. Then (65) means

$$Ae_i(x) \equiv \rho(x) \cdot e_i(x), \tag{66}$$

and consequently (64) becomes:

$$A(\varphi(x)e_i(x)) \equiv \rho(x)\varphi(x)e_i(x). \tag{67}$$

Summing (67) over $i = 1, \cdots, n$, and writing $f(x)$ for $\varphi(x) \sum_1^n e_i(x)$, we obtain

$$Af(x) \equiv \rho(x)f(x), \tag{68}$$

if $f(x)$ is bounded and $= 0$ for $x \notin T_0^{(1)} + \cdots + T_0^{(n)}$, for any $n = 1, 2, \cdots$.

A is bounded; so $\| Af \|_S \leqq C \cdot \| f \|_S$ for a suitable $C > 0$. Put $M_n = (x; | \rho(x) | \geqq C + 1, x \in T_0^{(1)} + \cdots + T_0^{(n)})$, $n = 1, 2, \cdots$. We may use

$$f(x) \equiv \theta_{M_n}(x) \begin{cases} = 1 & \text{for } x \in M_n \\ = 0 & \text{otherwise} \end{cases}$$

in (68). Then we have:

$$\| A\theta_{M_n} \|_S^2 = \int_S | A\theta_{M_n}(x) |^2 d_\mu x = \int_{M_n} | \rho(x) |^2 d_\mu x$$

$$\geqq \int_{M_n} (C+1)^2 d_\mu x = (C+1)^2 \mu(M_n),$$

$$\| \theta_{M_n} \|_S^2 = \int_S | \theta_{M_n}(x) |^2 d_\mu x = \int_{M_n} d_\mu x = \mu(M_n).$$

So $\| A\theta_{M_n} \|_S \leqq C \| \theta_{M_n} \|_S$ gives $\mu(M_n) = 0$. Consequently for $M = \sum_1^\infty M_n = (x; | \rho(x) | \geqq C + 1)$ also $\mu(M) = 0$. Hence we have (with the usual exception)

$$| \rho(x) | \leqq C + 1. \tag{69}$$

So $\rho(x)$ is bounded, and we may form $L_{\rho(x)}$. And (68) becomes:

$$Af(x) \equiv L_{\rho(x)} f(x) \tag{70}$$

for all $f(x)$ of (68). But these $f(x)$ form an everywhere dense set in $\mathfrak{H}_S$. Indeed, if any $f \in \mathfrak{H}_S$ is given, then define

$$f_n(x) \begin{cases} = f(x) & \text{for } |f(x)| \leqq n, x \in T_0^{(1)} + \cdots + T_0^{(n)} \\ = 0 & \text{otherwise} \end{cases}.$$

Then all $f_n(x)$, $n = 1, 2, \cdots$, satisfy the requirements stated after (68), and clearly $\lim_{n \to \infty} \| f_n - f \|_S = 0$.

As A, $L_{\rho(x)}$ are both bounded, this means that (70) extends by continuity to all $f \in \mathfrak{H}_S$. Hence $A = L_{\rho(x)} \in \mathfrak{L}$. So we have established $\mathfrak{L}' \subset \mathfrak{L}$ too, thus completing the proof.

LEMMA 3.5.4. *$L_{\varphi(x)} = L_{\psi(x)} U_\alpha$ holds if and only if either $\alpha = 1$, $\varphi(x) \equiv \psi(x)$ (except for an x-set of μ-measure 0), or $\alpha \neq 1$, $\varphi(x) \equiv \psi(x) \equiv 0$ (except for an x-set of μ-measure 0).*

PROOF: *Sufficiency*: Obvious.

Necessity: Assume $L_{\varphi(x)} = L_{\psi(x)} U_\alpha$. That is

$$\varphi(x) f(x) \equiv \psi(x) \left[\frac{d_{\mu_\alpha}}{d_\mu}(x) \right]^{\frac{1}{2}} f(x\alpha), \tag{71}$$

whenever $\int_S | f(x) |^2 d_\mu x$ is finite, (71), as well as all subsequent relations in this proof, holds with the exception of an x-set of μ-measure 0.

Consider the $T^{(1)}$, $T^{(2)}$, $\cdots$ of (iii) in Definition 3.2.4. We may put

$$f(x) \equiv \theta_{T^{(1)}+\cdots+T^{(n)}}(x) \begin{cases} = 1 & \text{for } x \in T^{(1)} + \cdots + T^{(n)} \\ = 0 & \text{otherwise} \end{cases}$$

in (71), since $\mu(T^{(1)} + \cdots + T^{(n)})$ is finite. Then (71) implies

$$\varphi(x) \equiv \psi(x) \left[\frac{d_{\mu_\alpha}}{d_\mu}(x)\right]^{\frac{1}{2}} \tag{72}$$

for those x for which $x, x\alpha \in T^{(1)} + \cdots + T^{(n)}$. And summing over $n = 1, 2, \cdots$ gives (72) for all x.

(72) settles the case $\alpha = \mathbf{1}$, owing to (ii) in Lemma 3.3.1. Assume therefore $\alpha \neq \mathbf{1}$. Substituting $\psi(x)$ from (72) into (71) gives $\varphi(x)f(x) \equiv \varphi(x)f(x\alpha)$, that is

$$\int_{\mathfrak{S}} |\varphi(x)|^2 |f(x) - f(x\alpha)|^2 d_\mu x = 0. \tag{73}$$

Now consider the system Δ of Definition 3.2.3. We have $\Delta = (M_1, M_2, \cdots)$ by (iii) eod. We may put

$$f(x) \equiv \theta_{T^{(i)} \cdot M_j}(x) \begin{cases} = 1 & \text{for } x \in T^{(i)} \cdot M_j \\ = 0 & \text{otherwise} \end{cases},$$

$i, j = 1, 2, \cdots$, since $\mu(T^{(i)} \cdot M_j) \leqq \mu(T^{(i)})$ is finite. Now clearly

$$|f(x) - f(x\alpha)| \begin{cases} = 0 & \text{if neither or both } x, x\alpha \in T^{(i)} \cdot M_j \\ = 1 & \text{otherwise} \end{cases}$$

and the latter x-set is

$$K_{ij} = ((T^{(i)} \cdot M_j) + (T^{(i)} \cdot M_j)\alpha^{-1}) - ((T^{(i)} \cdot M_j) \cdot (T^{(i)} \cdot M_j)\alpha^{-1}).$$

Hence (73) becomes

$$\int_{K_{ij}} |\varphi(x)|^2 d_\mu x = 0. \tag{74}$$

Consider now $K = \sum_{i,j=1}^{\infty} K_{ij}$. If $x = x\alpha$, then never $x \in K_{ij}$, and hence not $x \in K$. If $x \neq x\alpha$ then we know from (ii) in Definition 3.2.3 that for some $j = 1, 2, \cdots$ one, but not the other of x, $x\alpha$ is $\in M_j$. The one $\in M_j$ is also $\in T^{(i)}$ for some $i = 1, 2, \cdots$, hence one, but not the other of x, $x\alpha$ is $\in T^{(i)} \cdot M_j$. This means $x \in K_{ij}$. Consequently $x \in K$. So we see

$$K = (x; x \neq x\alpha). \tag{75}$$

Since $\mathfrak{G}$ is free, and as $\alpha \neq \mathbf{1}$, so we have by Definition 3.3.4, $x \neq x\alpha$, except for an x-set of μ-measure 0. That is

$$\mu(S - K) = 0. \tag{76}$$

Summing over $i, j = 1, 2, \cdots$ converts (74) into $\int_K |\varphi(x)|^2 d_\mu x = 0$, and so, by (76) $\int_S |\varphi(x)|^2 d_\mu x = 0$. That is

$$\varphi(x) \equiv 0. \tag{77}$$

So we have for $\alpha \neq \mathbf{1}$ by (72), (77), $\varphi(x) \equiv \psi(x) \equiv 0$, and the proof is completed.

LEMMA 3.5.5. *Let $\mathfrak{L}$ be as in Lemma 3.5.3, and let $\mathfrak{U}$ be the set of all U_α, $\alpha \in \mathfrak{G}$. Then $\mathfrak{L}\cdot\mathfrak{U}' = (aI)$ (a any complex number).*

PROOF: aI belongs obviously to $\mathfrak{U}'$, and as $aI = L_a \in \mathfrak{L}$, so $aI \in \mathfrak{L}\cdot\mathfrak{U}'$, $\mathfrak{L}\cdot\mathfrak{U}' \supset (aI)$. So we need only to prove $\mathfrak{L}\cdot\mathfrak{U}' \subset (aI)$.

Assume therefore $A \in \mathfrak{L}\cdot\mathfrak{U}'$. Then $A \in \mathfrak{U}'$ so for every $\alpha \in \mathfrak{G}$ $U_\alpha A = AU_\alpha$, $U_\alpha A U_\alpha^{-1} = A$. Also $A \in \mathfrak{L}$, $A = L_{\varphi(x)}$, $U_\alpha A U_\alpha^{-1} = L_{\varphi(x\alpha)}$, hence $L_{\varphi(x\alpha)} = L_{\varphi(x)}$. By Lemma 3.5.4 this means $\varphi(x\alpha) \equiv \varphi(x)$, except for an x-set of μ-measure 0. Since $\mathfrak{G}$ is ergodic, this implies by Lemma 3.3.2 that $\varphi(x) \equiv a$ (constant!) except for an x-set of μ-measure 0. Hence $A = L_a = aI$. Thus $\mathfrak{L}\cdot\mathfrak{U} \subset (aI)$, completing the proof.

3.6. We now pass to constructions in $\mathfrak{H}_S^{\mathfrak{G}}$.

LEMMA 3.6.1. *Define the operators*

(i) $\bar{U}_{\alpha_0}\mathcal{F}(x, \alpha) \equiv \left[\frac{d_\mu \alpha_0}{d_\mu}(x)\right]^{\frac{1}{2}} \mathcal{F}(x\alpha_0, \alpha\alpha_0)$

(ii) $\bar{V}_{\alpha_0}\mathcal{F}(x, \alpha) \equiv \mathcal{F}(x, \alpha_0^{-1}\alpha)$

for any $\alpha_0 \in \mathfrak{G}$.

(iii) $\bar{W}\mathcal{F}(x, \alpha) \equiv \left[\frac{d_\mu \alpha^{-1}}{d_\mu}(x)\right]^{\frac{1}{2}} \mathcal{F}(x\alpha^{-1}, \alpha^{-1})$.

(iv) $\bar{L}_{\varphi(x)}\mathcal{F}(x, \alpha) \equiv \varphi(x)\mathcal{F}(x, \alpha)$

(v) $\bar{M}_{\varphi(x)}\mathcal{F}(x, \alpha) \equiv \varphi(x\alpha^{-1})\mathcal{F}(x, \alpha)$

for any bounded and Γ-measurable $\varphi(x)$ which is defined for all $x \in S$.

These operators are bounded, $\bar{U}_{\alpha_0}$, $\bar{V}_{\alpha_0}$, $\bar{W}$ are even unitary.

PROOF: $\bar{U}_{\alpha_0}$, $\bar{V}_{\alpha_0}$, $\bar{W}$ are bounded and unitary: By (iii) in Lemma 3.3.1

$$\sum_{\alpha \in \mathfrak{G}} \int_S |\bar{U}_{\alpha_0}\mathcal{F}(x, \alpha)|^2 d_\mu x = \sum_{\alpha \in \mathfrak{G}} \int_S |\mathcal{F}(x\alpha_0, \alpha\alpha_0)|^2 \frac{d_\mu \alpha_0}{d_\mu}(x)\, d_\mu x$$

$$= \sum_{\alpha \in \mathfrak{G}} \int_S |\mathcal{F}(x, \alpha\alpha_0)|^2 d_\mu x = \sum_{\alpha \in \mathfrak{G}} \int_S |\mathcal{F}(x, \alpha)|^2 d_\mu x,$$

and

$$\sum_{\alpha \in \mathfrak{G}} \int_S |\bar{W}\mathcal{F}(x, \alpha)|^2 d_\mu x = \sum_{\alpha \in \mathfrak{G}} \int_S |\mathcal{F}(x\alpha^{-1}, \alpha^{-1})|^2 \frac{d_\mu \alpha^{-1}}{d}(x)\, d_\mu x$$

$$= \sum_{\alpha \in \mathfrak{G}} \int_S |\mathcal{F}(x, \alpha^{-1})|^2 d_\mu x = \sum_{\alpha \in \mathfrak{G}} \int_S |\mathcal{F}(x, \alpha)|^2 d_\mu x.$$

And obviously

$$\sum_{\alpha \in \mathfrak{G}} \int_S |\bar{V}_{\alpha_0} \mathcal{F}(x, \alpha)|^2 d_\mu x = \sum_{\alpha \in \mathfrak{G}} \int_S |\mathcal{F}(x, \alpha_0^{-1}\alpha)|^2 d_\mu x = \sum_{\alpha \in \mathfrak{G}} \int_S |\mathcal{F}(x, \alpha)|^2 d_\mu x.$$

So $\mathcal{F} \in \mathfrak{H}_S^{\mathfrak{G}}$ implies $\bar{U}_{\alpha_0}\mathcal{F}$, $\bar{V}_{\alpha_0}\mathcal{F}$, $\overline{W}\mathcal{F} \in \mathfrak{H}_S^{\mathfrak{G}}$, and $\|\bar{U}_{\alpha_0}\mathcal{F}\|_S^{\mathfrak{G}} = \|\bar{V}_{\alpha_0}\mathcal{F}\|_S^{\mathfrak{G}} = \|\overline{W}\mathcal{F}\|_S^{\mathfrak{G}} = \|\mathcal{F}\|_S^{\mathfrak{G}}$.

$\bar{L}_{\varphi(x)}$, $\bar{M}_{\varphi(x)}$ are bounded: We have $|\varphi(x)| \leqq C$ for all $x \in S$ and a suitable $C > 0$. Hence

$$\sum_{\alpha \in \mathfrak{G}} \int_S |\bar{L}_{\varphi(x)} \mathcal{F}(x, \alpha)|^2 d_\mu x = \sum_{\alpha \in \mathfrak{G}} \int_S |\varphi(x)|^2 |\mathcal{F}(x, \alpha)|^2 d_\mu x$$

$$\leqq C^2 \sum_{\alpha \in \mathfrak{G}} \int_S |\mathcal{F}(x, \alpha)|^2 d_\mu x,$$

and

$$\sum_{\alpha \in \mathfrak{G}} \int_S |\bar{M}_{\varphi(x)} \mathcal{F}(x, \alpha)|^2 d_\mu x = \sum_{\alpha \in \mathfrak{G}} \int_S |\varphi(x\alpha^{-1})|^2 |\mathcal{F}(x, \alpha)|^2 d_\mu x$$

$$\leqq C^2 \sum_{\alpha \in \mathfrak{G}} \int_S |\mathcal{F}(x, \alpha)|^2 d_\mu x.$$

So $\mathcal{F} \in \mathfrak{H}_S^{\mathfrak{G}}$ implies $\bar{L}_{\varphi(x)}\mathcal{F}$, $\bar{M}_{\varphi(x)}\mathcal{F} \in \mathfrak{H}_S^{\mathfrak{G}}$, and $\|\bar{L}_{\varphi(x)}\mathcal{F}\|_S^{\mathfrak{G}}$, $\|\bar{M}_{\varphi(x)}\mathcal{F}\|_S^{\mathfrak{G}} \leqq C\|\mathcal{F}\|_S^{\mathfrak{G}}$.

LEMMA 3.6.2. (i) $\bar{U}_1 = I$. (i)′ $\bar{V}_1 = I$.
(ii) $\bar{U}_{\alpha_0^{-1}} = \bar{U}_{\alpha_0}^{-1} = U_{\alpha_0}^*$. (ii)′ $\bar{V}_{\alpha_0^{-1}} = \bar{V}_{\alpha_0}^{-1} = \bar{V}_{\alpha_0}^*$.
(iii) $\bar{U}_{\alpha_0}\bar{U}_{\beta_0} = \bar{U}_{\alpha_0\beta_0}$. (iii)″ $\bar{V}_{\alpha_0}\bar{V}_{\beta_0} = \bar{V}_{\alpha_0\beta_0}$.
(iv) $\overline{W} = \overline{W}^{-1} = \overline{W}^*$.
(v) $\overline{W}\bar{U}_{\alpha_0}\overline{W} = \bar{V}_{\alpha_0}$.
(vi) $\overline{W}\bar{L}_{\varphi(x)}\overline{W} = \bar{M}_{\varphi(x)}$.

PROOF: We prove these statements in a somewhat changed order:

Ad (iv): We determine $\overline{W}^2$. One verifies immediately:

$\overline{W}^2\mathcal{F}(x, \alpha) \equiv \left[\frac{d_{\mu_{\alpha^{-1}}}}{d_\mu}(x)\frac{d_{\mu_\alpha}}{d_\mu}(x\alpha^{-1})\right]^{\frac{1}{2}} \mathcal{F}(x, \alpha)$. By (ii) in Lemma 3.3.1 (with α^{-1}, α in place of α, β) $\frac{d_{\mu_{\alpha^{-1}}}}{d_\mu}(x)\frac{d_{\mu_\alpha}}{d_\mu}(x\alpha^{-1}) \equiv \frac{d_{\mu_1}}{d_\mu}(x) \equiv 1$, so $\overline{W}^2\mathcal{F}(x, \alpha) \equiv \mathcal{F}(x, \alpha)$, that is: $\overline{W}^2 = I$.

Hence $\overline{W} = \overline{W}^{-1}$ and since $\overline{W}$ is unitary, so $\overline{W}^{-1} = \overline{W}^*$.

Ad (v): Owing to (iv) we may as well prove $\overline{W}\bar{V}_{\alpha_0}\overline{W} = \bar{U}_{\alpha_0}$. Now $\overline{W}\bar{V}_{\alpha_0}\overline{W}\mathcal{F}(x, \alpha) \equiv \left[\frac{d_{\mu_{\alpha^{-1}}}}{d_\mu}(x)\frac{d_{\mu_{\alpha\alpha_0}}}{d_\mu}(x\alpha^{-1})\right]^{\frac{1}{2}} \mathcal{F}(x\alpha_0, \alpha\alpha_0)$. By (iii) in Lemma 3.3.1 (with α^{-1}, $\alpha\alpha_0$ in place of α, β)

$\frac{d_{\mu_{\alpha^{-1}}}}{d_\mu}(x)\frac{d_{\mu_{\alpha\alpha_0}}}{d_\mu}(x\alpha^{-1}) \equiv \frac{d_{\mu_{\alpha_0}}}{d_\mu}(x)$, so $\overline{W}\bar{V}_{\alpha_0}\overline{W}\mathcal{F}(x,\alpha)$

$$\equiv \left[\frac{d_{\mu_{\alpha_0}}}{d_\mu}(x)\right]^{\frac{1}{2}}\mathcal{F}(x\alpha_0,\alpha\alpha_0) \equiv \bar{U}_{\alpha_0}\mathcal{F}(x,\alpha),$$

that is: $\overline{W}\bar{V}_{\alpha_0}\overline{W} = \bar{U}_{\alpha_0}$.

Ad (vi): Owing to (iv) we may as well prove $\overline{W}\bar{L}_{\varphi(x)} = \bar{M}_{\varphi(x)}\overline{W}$. This, however, is obviously true, since $\overline{W}\bar{L}_{\varphi(x)}\mathcal{F}(x,\alpha)$, $\bar{M}_{\varphi(x)}\overline{W}\mathcal{F}(x,\alpha)$ are both $\equiv \left[\frac{d_{\mu_{\alpha^{-1}}}}{d_\mu}(x)\right]^{\frac{1}{2}}\varphi(x\alpha^{-1})\mathcal{F}(x\alpha^{-1},\alpha^{-1})$.

Ad (i)′, (iii)′: Obvious.

Ad (ii)′: (i)′, (iii)′ above give $\bar{V}_{\alpha_0^{-1}} = \bar{V}_{\alpha_0}^{-1}$. And since $\bar{V}_{\alpha_0}$ is unitary, so $\bar{V}_{\alpha_0}^{-1} = \bar{V}_{\alpha_0}^{*}$.

Ad (i)–(iii): Considering (iv), (v) above, they result from (i)′–(iii)′ above, respectively.

LEMMA 3.6.3. *Let $\mathcal{J}$ be the set of all $\bar{U}_{\alpha_0}$ and all $\bar{L}_{\varphi(x)}$, and $\mathcal{J}$ the set of all $\bar{V}_{\alpha_0}$ and all $\bar{M}_{\varphi(x)}$. (Cf. Lemma* 3.6.1.) *Form for each bounded operator $\bar{A}$ in $\mathfrak{H}_s^{\mathfrak{G}}$ the decomposition from* (1), *p.* 136, *Definition* 2.4.2: $\bar{A} \sim <A_{\alpha,\beta}>$, $\alpha, \beta \in \mathfrak{G}$, *where every $A_{\alpha,\beta}$ is a bounded operator in $\mathfrak{H}_s$. (Cf. the decomposition $\mathfrak{H}_s^{\mathfrak{G}} = \mathfrak{H}^{\mathfrak{G}} \otimes \mathfrak{H}_s$ from Lemma* 3.4.1. *As described there, we replace the indices $t, s = 1, 2, \cdots$ by indices $\alpha, \beta \in \mathfrak{G}$, as already the complete, normalized, orthogonal set φ_{α_0}, $\alpha_0 \in \mathfrak{G}$, cf. loc. cit. above, in $\mathfrak{H}^{\mathfrak{G}}$ has been indexed this way.) Then we have:*

(i) *$\bar{A} \in \mathcal{J}'$ if and only if $A_{\alpha,\beta} = L_{\chi_{\alpha\beta^{-1}}(x\beta^{-1})}$.*

(ii) *$\bar{A} \in \mathcal{J}'$ if and only if $A_{\alpha,\beta} = L_{\chi_{\alpha^{-1}\beta}(x)}U_{\beta^{-1}\alpha}$.*

Here $\chi_\gamma(x)$ must be a (complex valued) Γ-measurable, bounded function of x for every $\gamma \in \mathfrak{G}$.

We do not determine for which systems $\chi_\gamma(x)$, $\gamma \in \mathfrak{G}$, a bounded $\bar{A}$ satisfying (i) *or* (ii) *actually exist. This, however, is true:*

(iii) *The systems $\chi_\gamma(x)$, $\gamma \in \mathfrak{G}$, obtained from all $\bar{A} \in \mathcal{J}'$ by* (i), *and those obtained from all $\bar{A} \in \mathcal{J}'$ by* (ii), *are identical.*

(iv) *An $\bar{A}_1 \in \mathcal{J}'$ and an $\bar{A}_2 \in \mathcal{J}'$ which have the same $\chi_\gamma(x)$, $\gamma \in \mathfrak{G}$, by* (i) *or* (ii) *respectively, obtain from each other by the involutory automorphism $\mathcal{F} \rightleftarrows \overline{W}\mathcal{J}$ of $\mathfrak{H}_s^{\mathfrak{G}}$.*

PROOF: Remember the definition of the $A_{\alpha,\beta}$ ((1), p. 136, Definition 2.4.2, using our present notations): $\bar{A}<f(x)\varphi_{\alpha'}(\alpha);\ \alpha \in \mathfrak{G}> = <A_{\alpha',\beta}f(x);\ \beta \in \mathfrak{G}>$; that is (if we replace α', α by α, β so that now $x \in S$ and $\beta \in \mathfrak{G}$ are the variables, and $\alpha \in \mathfrak{G}$ is a parameter): $\bar{A}f(x)\varphi_\alpha(\beta) = A_{\alpha,\beta}f(x)$. Then the definitions of $\bar{U}_{\alpha_0}$, $\bar{L}_{\varphi(x)}$, $\overline{W}$ give: If $\bar{A} \sim <A_{\alpha,\beta}>$, $\alpha, \beta \in \mathfrak{G}$, then

$$\bar{U}_{\alpha_0}^{-1}\bar{A}\bar{U}_{\alpha_0} \sim <U_{\alpha_0}^{-1}A_{\alpha\alpha_0^{-1},\beta\alpha_0^{-1}}U_{\alpha_0}>, \qquad \alpha, \beta \in \mathfrak{G}, \tag{78}$$

$$\overline{W}\bar{A}\overline{W} \sim <U_\beta^{-1}A_{\alpha^{-1},\beta^{-1}}U_\alpha>, \qquad \alpha, \beta \in \mathfrak{G}, \tag{79}$$

$$\bar{L}_{\varphi(x)}\bar{A} \sim <L_{\varphi(x)}A_{\alpha,\beta}>, \qquad \alpha, \beta \in \mathfrak{G}, \tag{80}$$

$$\bar{A}\bar{L}_{\varphi(x)} \sim <A_{\alpha,\beta}L_{\varphi(x)}>, \qquad \alpha, \beta \in \mathfrak{G}. \tag{81}$$

We now proceed to prove the statements (i)–(iv) in a somewhat changed order:

Ad (i): $\bar{A} \in \mathcal{J}'$ means that $\bar{A}$ commutes with all $\overline{L}_{\varphi(x)}$, $\overline{L}^*_{\varphi(x)}$, $\bar{U}_{\alpha_0}$, $\bar{U}^*_{\alpha_0}$. Since $\overline{L}^*_{\varphi(x)} = \overline{L}_{\overline{\varphi(x)}}$, $\bar{U}^*_{\alpha_0} = \bar{U}_{\alpha_0^{-1}}$, it suffices to consider the $\overline{L}_{\varphi(x)}$, $\bar{U}_{\varphi(x)}$. That is

$$\overline{L}_{\varphi(x)}\bar{A} = \bar{A}\overline{L}_{\varphi(x)}, \tag{82}$$

$$\bar{U}^{-1}_{\alpha_0}\bar{A}\bar{U}_{\alpha_0} = \bar{A}. \tag{83}$$

(82) means owing to (80), (81), that all $A_{\alpha,\beta} \in \mathfrak{L}'$, that is, by Lemma 3.5.3, $A_{\alpha,\beta} \in \mathfrak{L}$. So

$$A_{\alpha,\beta} = L_{\omega_{\alpha,\beta}(x)}, \tag{84}$$

$\omega_{\alpha,\beta}(x)$ a (complex-valued) Γ-measurable, bounded function of x.

(83) means, owing to (78), that

$$A_{\alpha,\beta} = U^{-1}_{\alpha_0} A_{\alpha\alpha_0^{-1},\beta\alpha_0^{-1}} U_{\alpha_0}. \tag{85}$$

Substituting (84) into (85) transforms it into $L_{\omega_{\alpha,\beta}(x)} = L_{\omega_{\alpha\alpha_0^{-1},\beta\alpha_0^{-1}}(x\alpha_0^{-1})}$ that is

$$\omega_{\alpha\beta}(x) \equiv \omega_{\alpha\alpha_0^{-1},\beta\alpha_0^{-1}}(x\alpha_0^{-1}), \tag{86}$$

except for an x-set of μ-measure 0.

If we put $\alpha_0 = \beta$, and write $\chi_\gamma(x)$ for $\omega_{\gamma,1}(x)$, then (86) specializes to

$$\omega_{\alpha,\beta}(x) \equiv \chi_{\alpha\beta^{-1}}(x\beta^{-1}), \tag{87}$$

except for an x-set of μ-measure 0.

Again (87) obviously implies (86).

So $\bar{A} \in \mathcal{J}'$ is equivalent to (84), (87), that is, to

$$A_{\alpha,\beta} = L_{\chi_{\alpha\beta^{-1}}(x\beta^{-1})}. \tag{88}$$

This completes the proof of (i).

Ad (ii), (iv): Since $\overline{W}$ is unitary, $\mathcal{F} \rightleftarrows \overline{W}\mathcal{F}$ is an automorphism of $\mathfrak{H}^{\mathfrak{S}}_s$, since $\overline{W} = \overline{W}^{-1}$, $\overline{W}^2 = I$ (by (iv) in Lemma 3.6.2), this automorphism is involutory. This automorphism carries an operator $\bar{A}$ in $\mathfrak{H}^{\mathfrak{S}}_s$ into the operator $\overline{W}\bar{A}\overline{W}$. Hence it carries $\bar{U}_{\alpha_0}$, $\overline{L}_{\varphi(x)}$ into $\bar{V}_{\alpha_0}$, $M_{\varphi(x)}$ (by (iv)–(vi) in Lemma 3.6.2), and so $\mathcal{J}$ into $\mathcal{J}$. Hence it carries $\mathcal{J}'$ into $\mathcal{J}'$.

So the general $\bar{A}_2 \in \mathcal{J}'$ obtains by forming $\bar{A}_2 = \overline{W}\bar{A}_1\overline{W}^{-1}$ for the general $\bar{A}_1 \in \mathcal{J}'$. Or (by (iv) in Lemma 3.6.2) equivalently by forming $\overline{W}\bar{A}_1\overline{W}$. Now (88) gives the $A_{\alpha,\beta}$ of the general $\bar{A}_1 \in \mathcal{J}'$, hence (79) permits us to find the $A_{\alpha,\beta}$ of the general $\bar{A}_2 \in \mathcal{J}'$. We obtain in this way for the latter $A_{\alpha,\beta}$:

$$A_{\alpha,\beta} = U^{-1}_{\beta} L_{\chi_{\alpha^{-1}\beta}(x\beta)} U_{\alpha}. \tag{89}$$

But since $U_\alpha = U_\beta U_{\beta^{-1}\alpha}$, $L_{\chi_\alpha{}^{-1}\beta(x\beta)}U_\beta = U_\beta L_{\chi_\alpha{}^{-1}\beta(x)}$, (89) can also be written

$$A_{\alpha,\beta} = L_{\chi_\alpha{}^{-1}\beta(x)}U_{\beta^{-1}\alpha}. \tag{90}$$

This completes the proof of (ii).

And we have also established all properties of $\mathcal{F} \rightleftarrows \overline{\mathcal{W}}\mathcal{F}$ claimed by (iv).

Ad (iii): Immediate by (iv).

REMARK: We have also seen (in the proof of (ii), (iv) above), that $\mathcal{F} \rightleftarrows \overline{\mathcal{W}}\mathcal{F}$ carries $\mathcal{I}$ into $\mathcal{J}$. And since it is involutory (cf. (iv) above), it also carries $\mathcal{J}$ into $\mathcal{I}$.

DEFINITION 3.6.1. (i) *We denote the set of all systems* $\chi_\gamma(x)$, $\gamma \in \mathfrak{G}$, *which occur in* (i) *or in* (ii) *in Lemma* 3.6.3 *(both are the same, cf.* (iii) *eod.) by* X_0. (ii) *We denote the relation between an* $\bar{A} \in \mathcal{I}'$ *or an* $\bar{A} \in \mathcal{J}'$ *and a system* $\chi_\gamma(x)$, $\gamma \in \mathfrak{G}$, *from* X_0, *as described in* (i) *or in* (ii) *in Lemma* 3.6.3, *respectively, in both cases by*

$$\bar{A} \approx \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}.$$

LEMMA 3.6.4. $\mathcal{R}(\mathcal{I}) = \mathcal{J}'$, $\mathcal{R}(\mathcal{J}) = \mathcal{I}'$.

PROOF: Since the involutory automorphism $\mathcal{F} \rightleftarrows \overline{\mathcal{W}}\mathcal{F}$ of $\mathfrak{S}_s^{\mathfrak{G}}$ interchanges $\mathcal{I}$ and $\mathcal{J}$ (cf. (iv) in Lemma 3.6.3 and the remark after this lemma), and consequently also $\mathcal{R}(\mathcal{I})$, $\mathcal{I}'$ and $\mathcal{R}(\mathcal{J})$, $\mathcal{J}'$, it suffices to prove $\mathcal{R}(\mathcal{I}) = \mathcal{J}'$.

Proof of $\mathcal{R}(\mathcal{I}) \subset \mathcal{J}'$: One verifies immediately, for $\bar{A} = \bar{U}_{\alpha_0}$

that $A_{\alpha,\beta}\begin{cases} = U_{\alpha_0} \text{ for } \alpha = \beta\alpha_0 \\ = 0 \quad \text{otherwise} \end{cases}$, and for $\bar{A} = \bar{L}_{\varphi(x)}$ that

$A_{\alpha,\beta}\begin{cases} = L_{\varphi(x)} \text{ for } \alpha = \beta \\ = 0 \quad \text{otherwise} \end{cases}$. So we have $A_{\alpha,\beta} = L_{\chi_\alpha{}^{-1}\beta(x)}U_{\beta^{-1}\alpha}$,

where $\chi_\gamma(x)\begin{cases} = 1 \text{ for } \gamma = \alpha_0^{-1} \\ = 0 \text{ otherwise} \end{cases}$, or $\chi_\gamma(x)\begin{cases} = \varphi(x) \text{ for } \gamma = 1 \\ = 0 \quad \text{otherwise} \end{cases}$,

respectively. So $\bar{A} \in \mathcal{J}'$ by (ii) in Lemma 3.6.3.

We have thus proved $\mathcal{I} \subset \mathcal{J}'$. Since $\mathcal{J}'$ is a ring, this implies $\mathcal{R}(\mathcal{I}) \subset \mathcal{J}'$.

Proof of $\mathcal{R}(\mathcal{I}) \supset \mathcal{J}'$: Consider any $\bar{A} \in \mathcal{I}'$, $\bar{B} \in \mathcal{J}'$, then we have by Lemma 3.6.3, $A_{\alpha,\beta} = L_{\chi_{\alpha\beta^{-1}(x\beta^{-1})}}$, $B_{\alpha,\beta} = L_{\omega_\alpha{}^{-1}\beta(x)}U_{\beta^{-1}\alpha}$. Put $\bar{C} = \bar{A}\bar{B}$, $\bar{D} = \bar{B}\bar{A}$, then (**1**), pp. 136–137, (iv) in Lemma 2.4.3, gives:

$$C_{\alpha,\beta} = \sum_{\gamma \in \mathfrak{G}} A_{\gamma,\beta}B_{\alpha,\gamma} = \sum_{\gamma \in \mathfrak{G}} L_{\chi_{\gamma\beta^{-1}(x\beta^{-1})}}L_{\omega_\alpha{}^{-1}\gamma(x)}U_{\gamma^{-1}\alpha}$$

$$= \sum_{\gamma \in \mathfrak{G}} L_{\chi_{\gamma\beta^{-1}(x\beta^{-1})}\,\omega_\alpha{}^{-1}\gamma(x)}U_{\gamma^{-1}\alpha}$$

$$= \sum_{\gamma \in \mathfrak{G}} L_{\chi_{\alpha\gamma^{-1}\beta^{-1}(x\beta^{-1})}\,\omega_\gamma{}^{-1}(x)}U_\gamma$$

(we replaced the summation index γ by $\alpha\gamma^{-1}$), and

$$\begin{aligned} D_{\alpha,\beta} &= \sum_{\gamma \in \mathfrak{G}} B_{\gamma,\beta} A_{\alpha,\gamma} = \sum_{\gamma \in \mathfrak{G}} L_{\omega_\gamma{}^{-1}\beta(x)} U_{\beta^{-1}\gamma} L_{\chi_{\alpha\gamma^{-1}}(x\gamma^{-1})} \\ &= \sum_{\gamma \in \mathfrak{G}} L_{\omega_\gamma{}^{-1}\beta(x)} L_{\chi_{\alpha\gamma^{-1}}(x\beta^{-1})} U_{\beta^{-1}\gamma} \\ &= \sum_{\gamma \in \mathfrak{G}} L_{\chi_{\alpha\gamma^{-1}}(x\beta^{-1})\,\omega_\gamma{}^{-1}\beta(x)} U_{\beta^{-1}\gamma} \\ &= \sum_{\gamma \in \mathfrak{G}} L_{\chi_{\alpha\gamma^{-1}\beta^{-1}}(x\beta^{-1})\,\omega_\gamma{}^{-1}(x)} U_\gamma \end{aligned}$$

(we replaced the summation index γ by $\beta\gamma$). Remembering that the order of summation does not matter in the above $\sum_{\gamma \in \mathfrak{G}}$ (cf. loc. cit. above), we see from the above equation that $C_{\alpha,\beta} = D_{\alpha,\beta}$ for all $\alpha, \beta \in \mathfrak{G}$. That is: $\bar{C} = \bar{D}$, or in other words: $\bar{A}\bar{B} = \bar{B}\bar{A}$. Since $\mathfrak{J}'$ is a ring, and $\bar{A} \in \mathfrak{J}'$ arbitrary, this means that $\bar{B} \in \mathfrak{J}''$. And as $\bar{B} \in \mathfrak{J}'$ was arbitrary, so we have $\mathfrak{J}' \subset \mathfrak{J}''$.

As $I \in \mathfrak{J}$ ($I = \bar{U}_1$), so $\mathfrak{R}(\mathfrak{J}) = \mathfrak{J}''$ by (**5**), p. 397, Theorem 7. Hence we have $\mathfrak{J}' \subset \mathfrak{R}(\mathfrak{J})$.

The proof is thereby completed.

LEMMA 3.6.5. *Put* $\mathfrak{M} = \mathfrak{R}(\mathfrak{J}) = \mathfrak{J}'$. *Then* $\mathfrak{M}' = \mathfrak{R}(\mathfrak{J}) = \mathfrak{J}'$, *and* $\mathfrak{M}, \mathfrak{M}'$ *are factors* (*in the sense of* (**1**), *p.* 138, *Definition* 3.1.2).

PROOF: Clearly $\mathfrak{M}' = \mathfrak{R}(\mathfrak{J})' = \mathfrak{J}'$, this and Lemma 3.6.4 prove all our equations. It remains for us to prove that $\mathfrak{M}$, $\mathfrak{M}'$ are factors, that is that $\bar{A} \in \mathfrak{M} \cdot \mathfrak{M}'$ implies $\bar{A} = a\mathfrak{J}$.

Consider an $\bar{A} \in \mathfrak{M} \cdot \mathfrak{M}'$. Then $\bar{A} \in \mathfrak{M} = \mathfrak{J}'$ and $\bar{A} \in \mathfrak{M}' = \mathfrak{J}'$, hence Lemma 3.6.3 gives that

$$A_{\alpha,\beta} = L_{\chi_{\alpha\beta^{-1}}(x\beta^{-1})} \tag{91}$$

and

$$A_{\alpha,\beta} = L_{\omega_\alpha{}^{-1}\beta(x)} U_{\beta^{-1}\alpha} \tag{92}$$

simultaneously. Comparing (91), (92), and using Lemma 3.5.4 gives:

For $\beta^{-1}\alpha \neq 1$, that is $\alpha \neq \beta$: $\chi_{\alpha\beta^{-1}}(x\beta^{-1}) \equiv \omega_\alpha{}^{-1}{}_\beta(x) \equiv 0$, except for an x-set of μ-measure 0. That is: In this case (91) gives $A_{\alpha,\beta} = 0$.

For $\beta^{-1}\alpha = 1$, that is $\alpha = \beta$: $\chi_1(x\beta^{-1}) \equiv \omega(x)$, except for an x-set of μ-measure 0. Comparing herein $\beta = \mathbf{1}$ with the general $\beta \in \mathfrak{G}$ gives $\chi_1(x\beta^{-1}) \equiv \chi_1(x)$, with the same exceptions. Since $\mathfrak{G}$ is ergodic, this gives, by Lemma 3.3.2 (with $\chi_1(x)$, β^{-1} in place of $f(x)$, α): $\chi_1(x) \equiv a$, except for an x-set of μ-measure 0. Hence (91) gives $A_{\alpha,\alpha} = L_a = aI$.

So we see:

$$A_{\alpha,\beta} \begin{cases} = aI \text{ for } \alpha = \beta\,, \\ = 0 \quad \text{otherwise} \end{cases}$$

that is: $\bar{A} = aI$. Thus the proof is completed.

3.7. We derive now the basic rules of algebraical computation in the factors $\mathfrak{M}$, $\mathfrak{M}'$.

LEMMA 3.7.1. *Both in $\mathfrak{M}$ and in $\mathfrak{M}'$ the following rules of computation hold*: *If $\bar{A} \sim \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}$, $\bar{B} \sim \{\xi_\gamma(x); \gamma \in \mathfrak{G}\}$, $\bar{C} \sim \{\eta_\gamma(x); \gamma \in \mathfrak{G}\}$ (all of them $\in \mathfrak{M}$, or all of them $\in \mathfrak{M}'$!), then we have*:

(i) *If $\bar{C} = a\bar{A}$, then $\eta_\gamma(x) \equiv a\chi_\gamma(x)$.*

(ii) *If $\bar{C} = \bar{A}^*$, then $\eta_\gamma(x) \equiv \overline{\chi_{\gamma^{-1}}(x\gamma^{-1})}$.*

(iii) *If $\bar{C} = \bar{A} + \bar{B}$, then $\eta_\gamma(x) \equiv \chi_\gamma(x) + \xi_\gamma(x)$.*

(iv) *If $\bar{C} = \bar{A}\bar{B}$, then $\eta_\gamma(x) \equiv \sum_{\alpha \in \mathfrak{G}} \chi_\alpha(x)\xi_{\gamma\alpha^{-1}}(x\alpha^{-1})$. The $\sum_{\alpha \in \mathfrak{G}}$ converges "en mesure" (cf. the precise description of this notion at the end of the proof), irrespective of the order in which the $\alpha \in \mathfrak{G}$ are gone through.*

PROOF: Considering (iv) in Lemma 3.6.3, it suffices to consider the case where $\bar{A}, \bar{B}, \bar{C} \in \mathfrak{M}$. We prove the statements (i)–(iv) for this case in a somewhat changed order:

Ad (i), (iii): Immediate by (i), (iii) in (**1**), pp. 136–137, Lemma 2.4.3.

Ad (ii): This follows from (ii), loc. cit. above, by the following consideration: $C_{\alpha,\beta} = L_{\eta_{\alpha\beta^{-1}}(x\beta^{-1})}$, $C_{\alpha,\beta} = A^*_{\beta,\alpha} = L_{\overline{\chi_{\beta\alpha^{-1}}(x\alpha^{-1})}}$, $\eta_{\alpha\beta^{-1}}(x\beta^{-1}) \equiv \chi_{\beta\alpha^{-1}}(x\alpha^{-1})$, that is $\eta_\gamma(x) \equiv \chi_{\gamma^{-1}}(x\gamma^{-1})$.

Ad (iv): Apply (iv), loc. cit. above. We have $C_{\alpha,\beta} = L_{\eta_{\alpha\beta^{-1}}(x\beta^{-1})}$, $C_{\alpha,\beta} = \sum_{\gamma \in \mathfrak{G}} A_{\gamma,\beta}B_{\alpha,\gamma} = \sum_{\gamma \in \mathfrak{G}} L_{\chi_{\gamma\beta^{-1}}(x\beta^{-1})}L_{\xi_{\alpha\gamma^{-1}}(x\gamma^{-1})} = \sum_{\gamma \in \mathfrak{G}} L_{\chi_{\gamma\beta^{-1}}(x\beta^{-1})\xi_{\alpha\gamma^{-1}}(x\gamma^{-1})}$ in the sense of strong convergence, irrespective of the order of the $\gamma \in \mathfrak{G}$. So we have $L_{\eta_{\alpha\beta^{-1}}(x\beta^{-1})} = \sum_{\gamma \in \mathfrak{G}} L_{\chi_{\gamma\beta^{-1}}(x\beta^{-1})\xi_{\alpha\gamma^{-1}}(x\gamma^{-1})}$. Putting $\beta = 1$ and writing α, γ in place of γ, α this becomes:

$$L_{\eta_\gamma(x)} = \sum_{\alpha \in \mathfrak{G}} L_{\chi_\alpha(x)\xi_{\gamma\alpha^{-1}}(x\alpha^{-1})}. \tag{93}$$

in the sense of strong convergence, irrespective of the order of the $\alpha \in \mathfrak{G}$.

Consider now an $M \in \Gamma$ with a finite $\mu(M)$, and form

$$\theta_M(x) \begin{cases} = 1 \text{ for } x \in M \\ = 0 \text{ otherwise} \end{cases}.$$

Then $\theta_M \in \mathfrak{H}_S$, and so (93) implies: If $\bar{\alpha}_1, \bar{\alpha}_2, \cdots$ is an enumeration of $\mathfrak{G}$, then

$$\lim_{n\to\infty} \left\| L_{\eta_\gamma(x)}\theta_M - \sum_{i=1}^{n} L_{\chi_{\bar{\alpha}_i}(x)\xi_{\gamma\bar{\alpha}_i^{-1}}(x\bar{\alpha}_i^{-1})}\theta_M \right\|_S^2 = 0.$$

That is

$$\lim_{n\to\infty} \int_S \left| \eta_\gamma(x)\theta_M(x) - \sum_{i=1}^{n} \chi_{\bar{\alpha}_i}(x)\xi_{\gamma\bar{\alpha}_i^{-1}}(x\bar{\alpha}_i^{-1})\theta_M(x) \right|^2 d_\mu x = 0,$$

$$\lim_{n\to\infty} \int_M \left| \eta_\gamma(x) - \sum_{i=1}^{n} \chi_{\bar{\alpha}_i}(x)\xi_{\gamma\bar{\alpha}_i^{-1}}(x\bar{\alpha}_i^{-1}) \right|^2 d_\mu x = 0 \tag{94}$$

So we have for every $\epsilon > 0$ and $M \in \Gamma$ with a finite $\mu(M)$, owing to (94):

$$\lim_{n\to\infty} \mu\left(\left(x; \left|\eta_\gamma(x) - \sum_{i=1}^{n} \chi_{\bar\alpha_i}(x)\xi_{\gamma\bar\alpha_i^{-1}}(x\bar\alpha_i^{-1})\right| \geqq \epsilon\right)\cdot M\right) = 0. \tag{95}$$

That is: $\sum_{i=1}^{\infty} \chi_{\bar\alpha_i}(x)\xi_{\gamma\bar\alpha_i^{-1}}(x\bar\alpha_i^{-1})$ converges "en mesure" to $\eta_\gamma(x)$. As $\bar\alpha_1$, $\bar\alpha_2$, $\cdots$ was an arbitrary enumeration of $\mathfrak{G}$ this proves (iv). (Of course, if $\mathfrak{G}$ is finite, the convergence considerations are unnecessary, the finite sum $\sum_{\alpha\in\mathfrak{G}}$ being simply equal to $\eta_\gamma(x)$.)—

We sum up the results obtained so far.

THEOREM VIII. *Let S be a fixed set, Γ a separable Borel-system in S (Definitions* 3.2.1, 3.2.3), *μ a Γ-measure in S (Definition* 3.2.4) *and $\mathfrak{G}$ an S, Γ, μ-group (Definitions* 3.3.1–3.3.3; *hence $\mathfrak{G}$ is, in particular, finite or enumerably infinite), which is free and ergodic (Definitions* 3.3.4, 3.3.5). *Form the unitary (finite dimensional Euclidean or Hilbert) spaces $\mathfrak{H}_S$, $\mathfrak{H}^{\mathfrak{G}}$, $\mathfrak{H}_S^{\mathfrak{G}}$ (Definitions* 3.2.6, 3.4.1). *In $\mathfrak{H}_S^{\mathfrak{G}}$ form the operators $\bar U_{\alpha_0}$, $\bar V_{\alpha_0}$, $\overline{W}$, $\bar L_{\varphi(x)}$, $\bar M_{\varphi(x)}$ (Lemma* 3.6.1). *Let $\mathfrak{I}$ be the set of all $\bar U_{\alpha_0}$, $\bar L_{\varphi(x)}$ and $\mathfrak{J}$ the set of all $\bar V_{\alpha_0}$, $\bar M_{\varphi(x)}$. Then we have:*

$$\mathfrak{M} = \mathfrak{R}(\mathfrak{I}) = \mathfrak{J}' \text{ is a ring, its } ' \text{ is } \mathfrak{M}' = \mathfrak{R}(\mathfrak{J}) = \mathfrak{I}'.$$

$\mathcal{F} \rightleftarrows \overline{W}\mathcal{F}$ is an involutory automorphism of $\mathfrak{H}_S^{\mathfrak{G}}$ which interchanges $\mathfrak{I}$, $\mathfrak{R}(\mathfrak{I})$, $\mathfrak{I}'$, $\mathfrak{M}$ and $\mathfrak{J}$, $\mathfrak{R}(\mathfrak{J})$, $\mathfrak{J}'$, $\mathfrak{M}'$. $\mathfrak{M}$, $\mathfrak{M}'$ are factors. The general element $\bar A \in \mathfrak{M}$ or $\bar A \in \mathfrak{M}'$ has a representation $\bar A \approx \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}$ (Definition 3.6.1), *for which the rules of algebraical computation are known (Lemma* 3.7.1).

PROOF: This is merely a restatement of the essential results of various Definitions and Lemmas in §§3.2–3.7. Besides those mentioned in the text, cf. also Lemmas 3.6.4, 3.6.5, (iv) in Lemma 3.6.3, and finally (i), (ii) in Lemma 3.6.3.—

The results and notions enumerated in Theorem VIII will be made use of throughout what follows, without being referred to explicitly.

We conclude this section by two lemmata on certain special operators in $\mathfrak{M}$ and $\mathfrak{M}'$. (All -relations in both of them are to be understood to hold with the exception of x-sets of μ-measure 0 only.)

LEMMA 3.7.2. *Assume $\bar A \in \mathfrak{M}$ or $\mathfrak{M}'$, $\bar A \approx \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}$. Then we have:*

(i) *$\bar A$ is Hermitean if and only if $\chi_\gamma(x) \equiv \overline{\chi_{\gamma^{-1}}(x\gamma^{-1})}$ for all $\gamma \in \mathfrak{G}$. This implies that $\chi_1(x)$ is real.*

(ii) *If $\bar A$ is definite, then $\chi_1(x) \geqq 0$.*

(iii) *If $\bar A$ is definite and $\chi_1(x) \equiv 0$, then $\bar A = 0$.*

PROOF: Considering the isomorphism $\mathcal{F} \rightleftarrows \overline{W}\mathcal{F}$ of $\mathfrak{M}$, $\mathfrak{M}'$ (cf. in particular (iv) in Lemma 3.6.3), it suffices to consider the case $\bar A \in \mathfrak{M}'$.

Ad (i): $\bar A \approx \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}$, $\bar A^* \approx \{\overline{\chi_{\gamma^{-1}}(x\gamma^{-1})}; \gamma \in \mathfrak{G}\}$. $\bar A$ Hermitean means $\bar A = \bar A^*$, that is $\chi_\gamma(x) \equiv \overline{\chi_{\gamma^{-1}}(x\gamma^{-1})}$. For $\gamma = \mathbf{1}$ this becomes $\chi_1(x) \equiv \overline{\chi_1(x)}$, so $\chi_1(x)$ must be real.

Ad (ii): Consider an $f \in \mathfrak{H}_S$ and define

$$\mathcal{F}(x, \alpha) \begin{cases} = f(x) \text{ for } \alpha = \mathbf{1} \\ = 0 \quad \text{otherwise} \end{cases}.$$

Then $\mathcal{F} \in \mathfrak{H}_S^{\mathcal{G}}$, and as $\bar{A}$ is definite, $(\bar{A}\mathcal{F}, \mathcal{F})_S^{\mathcal{G}} \geqq 0$. Now one verifies easily that $(\bar{A}\mathcal{F}, \mathcal{F})_S^{\mathcal{G}} = \int_S \chi_1(x) \, |f(x)|^2 \, d_\mu x$. So we see:

$$\int_S \chi_1(x) \, |f(x)|^2 \, d_\mu x \geqq 0 \quad \text{for every} \quad f \in \mathfrak{H}_S. \tag{96}$$

If $M \in \Gamma$, $\mu(M)$ finite, we can put

$$f(x) \equiv \theta_M(x) \begin{cases} = 1 \text{ for } x \in M \\ = 0 \text{ otherwise} \end{cases},$$

then $f \in \mathfrak{H}_S$, and (96) becomes

$$\int_S \chi_1(x) \, d_\mu x \geqq 0. \tag{97}$$

Now $\chi_1(x)$ is real by (i) above. Consider the $T^{(1)}, T^{(2)}, \cdots$ of (iii) in Definition 3.2.4, and put $M = (x; \chi_1(x) < 0) \cdot T^{(i)}$. Then $\mu(M) \leqq \mu(T^{(i)})$, hence it is finite, so (97) holds. And as $\chi_1(x) < 0$ for all $x \in M$, so $\mu(M) = 0$. Thus $\chi_1(x) \geqq 0$ for all $x \in T^{(i)}$ (exceptions as above!). Summing over $i = 1, 2, \cdots$ gives $\chi_1(x) \geqq 0$, unrestrictedly.

Ad (iii): Consider an $\alpha_0 \in \mathcal{G}$, $\alpha_0 \neq \mathbf{1}$, and two $f, g \in \mathfrak{H}_S$. Define

$$\mathcal{F}(x, \alpha) \begin{cases} = f(x) \text{ for } \alpha = \mathbf{1} \\ = g(x) \text{ for } \alpha = \alpha_0 \\ = 0 \quad \text{otherwise} \end{cases}.$$

Then $\mathcal{F} \in \mathfrak{H}_S^{\mathcal{G}}$ and, as $\bar{A}$ is definite, $(\bar{A}\mathcal{F}, \mathcal{F})_S^{\mathcal{G}} \geqq 0$. Now one verifies easily, considering $\chi_1(x) \equiv 0$ that

$$(\bar{A}\mathcal{F}, \mathcal{F})_S^{\mathcal{G}} = \int_S \chi_{\alpha_0}(x) f(x) \overline{g(x)} \, d_\mu x + \int_S \chi_{\alpha_0^{-1}}(x\alpha_0^{-1}) g(x) \overline{f(x)} \, d_\mu(x). \tag{98}$$

Now by (i), $\chi_{\alpha_0}(x) \equiv \chi_{\alpha_0^{-1}}(x\alpha_0^{-1})$, so the second term in (98) is the complex conjugate of the first one. Hence (98) becomes

$$(\bar{A}\mathcal{F}, \mathcal{F})_S^{\mathcal{G}} = 2\Re \int_S \chi_{\alpha_0}(x) f(x) \overline{g(x)} \, d_\mu x. \tag{99}$$

By (99), $(A\mathcal{F}, \mathcal{F})_S^{\mathcal{G}} \geqq 0$ states that

$$\Re \int_S \chi_{\alpha_0}(x) f(x) \overline{g(x)} \, d_\mu x \geqq 0. \tag{100}$$

If $M \in \Gamma$, $\mu(M)$ finite, we can put

$$f(x)\begin{cases} = 1 \text{ for } x \in M \\ = 0 \text{ otherwise} \end{cases},$$

$$g(x)\begin{cases} = -\operatorname{sgn} \chi_1(x) \text{ for } x \in M^{12} \\ = 0 \qquad \text{otherwise} \end{cases},$$

then $f, g \in \mathfrak{H}_S$, and (100) becomes

$$\int_S |\chi_{\alpha_0}(x)|\, d_\mu x \leqq 0. \tag{101}$$

Thus $\chi_{\alpha_0}(x) \equiv 0$ for all $x \in M$ (exceptions as above!). Now with the $T^{(1)}$, $T^{(2)}, \cdots$ of (iii) in Definition 3.2.4 we may put $M = T^{(i)}$ and sum over $i = 1, 2, \cdots$; this gives $\chi_{\alpha_0}(x) \equiv 0$ unrestrictedly.

So we have $\chi_\alpha(x) \equiv 0$ for $\alpha \neq \mathbf{1}$. For $\alpha = \mathbf{1}$ it holds by assumption. Thus $\bar{A} = 0$.—

The statements which follow will be made for $\mathfrak{M}$ only, but the automorphism $\mathcal{F} \rightleftarrows \overline{W}\mathcal{F}$ in $\mathfrak{H}_S^{\mathfrak{G}}$ (cf. in particular (iv) in Lemma 3.6.3) extends them immediately to $\mathfrak{M}'$ too.

Definition 3.7.1. (i) *We denote the set of all* $\overline{L}_{\varphi(x)}$ *by* $\overline{\mathfrak{L}}$.

(ii) *For every* $M \in \Gamma$ *put* $\bar{E}(M) = \overline{L}_{\theta_M(x)}$, *where* $\theta_M(x)\begin{cases} = 1 \text{ for } x \in M \\ = 0 \text{ otherwise} \end{cases}$.

Lemma 3.7.3. (i) $\overline{\mathfrak{L}}$ *is an Abelian ring* $\subset \mathfrak{M}$, *we have*

$$\overline{\mathfrak{L}} = \mathfrak{M} \cdot \overline{\mathfrak{L}}'.$$

(ii) *The projections in* $\overline{\mathfrak{L}}$ *coincide with the* $\bar{E}(M)$.

Proof: Ad (i): It suffices to prove $\overline{\mathfrak{L}} = \mathfrak{M} \cdot \overline{\mathfrak{L}}'$: Since $\mathfrak{M} \cdot \overline{\mathfrak{L}}'$ is clearly a ring, it proves that $\overline{\mathfrak{L}}$ is also a ring, and implies $\overline{\mathfrak{L}} \subset \mathfrak{M}$, and also $\overline{\mathfrak{L}} \subset \overline{\mathfrak{L}}'$, that is, that $\overline{\mathfrak{L}}$ is Abelian.

Proof of $\overline{\mathfrak{L}} \subset \mathfrak{M} \cdot \overline{\mathfrak{L}}'$: $\overline{L}_{\varphi(x)} \in \mathcal{J} \subset \mathcal{R}(\mathcal{J}) = \mathfrak{M}$, so $\overline{\mathfrak{L}} \subset \mathfrak{M}$. Every $\overline{L}_{\varphi(x)}$ commutes with every $\overline{L}_{\psi(x)}$, and $\overline{L}^*_{\psi(x)} = \overline{L}_{\overline{\psi(x)}}$, hence it is $\in \overline{\mathfrak{L}}'$, so $\overline{\mathfrak{L}} \subset \overline{\mathfrak{L}}'$. Hence $\overline{\mathfrak{L}} \subset \mathfrak{M} \cdot \overline{\mathfrak{L}}'$.

Proof of $\overline{\mathfrak{L}} \supset \mathfrak{M} \cdot \overline{\mathfrak{L}}'$: Consider an $\bar{A} \in \mathfrak{M} \cdot \overline{\mathfrak{L}}'$. Since $\bar{A} \in \mathfrak{M}$, so $\bar{A} \approx \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}$. Since $\bar{A} \in \overline{\mathfrak{L}}'$, so $\bar{A}$ commutes with every $\overline{L}_{\varphi(x)}$. Now $\overline{L}_{\varphi(x)} \approx \{\varphi(x)\delta_\gamma; \gamma \in \mathfrak{G}\}$, where

$$\delta_\gamma\begin{cases} = 1 \text{ for } \gamma = \mathbf{1} \\ = 0 \text{ otherwise} \end{cases}.$$

---

[12] For any complex a $\operatorname{sgn} a\begin{cases} = \dfrac{a}{|a|} & \text{for } a \neq 0 \\ = 0 & \text{for } a = 0 \end{cases}$. So always $|\operatorname{sgn} a| \leqq 1$ and $a \overline{\operatorname{sgn} a} = |a|$.

So $\bar{L}_{\varphi(x)}\bar{A} \approx \{\varphi(x)\chi_\gamma(x);\ \gamma \in \mathfrak{G}\}$, $\bar{A}\bar{L}_{\varphi(x)} \approx \{\varphi(x\gamma^{-1})\chi_\gamma(x);\ \gamma \in \mathfrak{G}\}$. That is

$$\varphi(x)\chi_\gamma(x) \equiv \varphi(x\gamma^{-1})\chi_\gamma(x). \tag{102}$$

Put

$$\varphi(x) \equiv \theta_M(x)\left\{\begin{matrix}= 1 \text{ for } x \in M\\ = 0 \text{ otherwise}\end{matrix}\right|,$$

for an $M \in \Gamma$, then (102) becomes

$$\chi_\gamma(x) \equiv 0 \quad \text{for} \quad x \in M,\ x\gamma^{-1} \notin M, \quad \text{or} \quad x\gamma^{-1} \in M,\ x \notin M. \tag{103}$$

Now let M run over all elements of the finite or enumerably infinite system Δ in Definition 3.2.3; owing to (ii) eod. (103) becomes, if we sum over all $M \in \Delta$:

$$\chi_\gamma(x) \equiv 0 \tag{104}$$

if $x\gamma^{-1} \neq x$ that is, if $x \neq x\gamma$. Now if $\gamma \neq \mathbf{1}$, then since $\mathfrak{G}$ is free, this means that (104) holds without any further qualifications. So $\chi_\gamma(x) \equiv \chi_1(x)\delta_\gamma$ (δ_γ as above), and hence $\bar{A} = \bar{L}_{\chi_1(x)}$ (see above), and therefore $\bar{A} \in \bar{\mathfrak{L}}$.

Thus we have shown that $\mathfrak{M}\cdot\bar{\mathfrak{L}}' \subset \bar{\mathfrak{L}}$, and thereby completed the proof.

Ad (ii): That an $\bar{A} \in \bar{\mathfrak{L}}$ is a projection, means this: $\bar{A} = \bar{L}_{\varphi(x)}$ and $\bar{A} = \bar{A}^*$, $\bar{A} = \bar{A}^2$. But $\bar{A}^* = \bar{L}_{\overline{\varphi(x)}}$, $\bar{A}^2 = \bar{L}_{\varphi(x)^2}$, so it means $\varphi(x) \equiv \overline{\varphi(x)}$, $\varphi(x) \equiv \varphi(x)^2$. That is: $\varphi(x) \equiv 0, 1$. But these $\varphi(x)$ are precisely the

$$\varphi(x) \equiv \theta_M(x)\left\{\begin{matrix}= 1 \text{ for } x \in M\\ = 0 \text{ otherwise}\end{matrix}\right|$$

for all $M \in \Gamma$, and so the $\bar{A}$ are the $\bar{L}_{\theta_M(x)} = \bar{E}(M)$ for all $M \in \Gamma$.

Chapter IV. Determination of the Classes of $\mathfrak{M}$, $\mathfrak{M}'$

4.1. The essential distinction which has to be made when determining the classes of $\mathfrak{M}$, $\mathfrak{M}'$, is whether the S, Γ, μ-group $\mathfrak{G}$ is measurable or not. (Cf. Definition 3.3.6.)

Consider first the case when $\mathfrak{G}$ is measurable. This means that a Γ-measure ν as described in Definition 3.3.6 exists, or, equivalently, that a function $\kappa(x)$ as described in Lemma 3.3.3 exists.

In this case it is possible to replace μ by ν altogether: $\mathfrak{G}$ is an S, Γ, ν-group as well as an S, Γ, μ-group, and its being free and ergodic depends only on the sets of measure 0—which are the same for μ and for ν. $\mathfrak{H}_{S,\Gamma,\mu}$ and $\mathfrak{H}^{\mathfrak{G}}_{S,\Gamma,\mu}$ become $\mathfrak{H}_{S,\Gamma,\nu}$ and $\mathfrak{H}^{\mathfrak{G}}_{S,\Gamma,\nu}$ if we replace $f(x)$ and $\mathcal{F}(x, \alpha)$ by $(\kappa(x))^{\frac{1}{2}}f(x)$ and $(\kappa(x))^{\frac{1}{2}}\mathcal{F}(x, \alpha)$ respectively. That is: This replacement compensates for the changes brought about by substituting ν for μ in $\mathfrak{H}_S$ and $\mathfrak{H}^{\mathfrak{G}}_S$. And finally (ii) in Lemma 3.3.3 shows that this replacement has also the right effect in the definitions of $\bar{U}_{\alpha_0}$, $\bar{V}_{\alpha_0}$, $\bar{W}$, $\bar{L}_{\varphi(x)}$, $\bar{M}_{\varphi(x)}$.

In other words: If $\mathfrak{G}$ is measurable at all, then we may even assume that μ itself has the properties required for ν, especially (ii) in Definition 3.3.6. That is:

$$\mu(M) = \mu(M\alpha) = \mu_\alpha(M), \frac{d_{\mu_\alpha}}{d_\mu}(x) \equiv 1. \tag{105}$$

This means, however, that for a measurable $\mathfrak{G}$ our $\mathfrak{M}$, $\mathfrak{M}'$ coincide essentially with those of (**1**), pp. 192–204, and so the exhaustive discussion given loc. cit. disposes of our $\mathfrak{M}$, $\mathfrak{M}'$ also in this case. Indeed, it is for the sake of the not measurable $\mathfrak{G}$, to be discussed in §§4.3–4.4 below, that this work was undertaken. (They correspond to case (III_∞), cf. Theorem IX.)

We shall nevertheless, for the sake of completeness, discuss the measurable $\mathfrak{G}$ also, and that without bothering about replacing μ by ν. We shall, however, keep this discussion as short as possible, and not enter into questions which could also be answered now: Find criteria for normedness, expressions for $\overline{T}_{\mathfrak{M}}(\bar{A})$, $[[\bar{A}]]$, $\ll\bar{A}, \bar{B}\gg$, etc.

4.2. We assume throughout this section that $\mathfrak{G}$ is measurable, and make use of the (essentially unique) ν, $\kappa(x)$ of Definition 3.3.6 and Lemma 3.3.3.

All our results, throughout §§4.2–4.4, will be stated and proved for $\mathfrak{M}$ only: The automorphism $\mathcal{F} \rightleftarrows \overline{W}\mathcal{F}$ of $\mathfrak{H}_g^{\mathfrak{G}}$ (cf. in particular (iv) in Lemma 3.6.6) extends them immediately to $\mathfrak{M}'$ too.

DEFINITION 4.2.1. *If $\bar{A} \in \mathfrak{M}$ is definite, then we have for $\bar{A} \approx \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}$ (by* (ii) *in Lemma* 3.7.2*) $\chi_1(x) \geqq 0$. Hence we can certainly form the integral*

$$\int_g \chi_1(x)\kappa(x)\, d_\mu x.$$

Its numerical value is $\geqq 0$, but it may be finite or infinite. We denote this value by $T^0(\bar{A})$.

LEMMA 4.2.1. *We have for $\bar{A}$, $\bar{B}$, $\bar{C} \in \mathfrak{M}$, $\bar{A}$, $\bar{B}$ definite:*

(i) $0 \leqq T^0(\bar{A}) \leqq +\infty$.

(ii) $T^0(\bar{A}) > 0$ if $\bar{A} \neq 0$.

(iii) $T^0(\bar{A} + \bar{B}) = T^0(\bar{A}) + T^0(\bar{B})$.

(iv) $\bar{C} \approx \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}$. *$\bar{C}^*\bar{C}$, $\bar{C}\bar{C}^*$ are both $\in \mathfrak{M}$ and definite, and*

$$T^0(\bar{C}^*\bar{C}) = T^0(\bar{C}\bar{C}^*) = \sum_{\gamma \in \mathfrak{G}} \int_g |\chi_\gamma(x)|^2 \kappa(x)\, d_\mu x.$$

(*Since all terms in this $\sum_{\gamma \in \mathfrak{G}}$ are $\geqq 0$, it is clear what it means.*)

PROOF: Ad (i): Immediate by (ii) in Lemma 3.7.2.

Ad (ii): Immediate by (ii), (iii) in Lemma 3.7.2

Ad (iii): Obvious

Ad (iv): $\bar{C}^*\bar{C}, \bar{C}\bar{C}^* \in \mathfrak{M}$ along with $\bar{C} \in \mathfrak{M}$, their definiteness is obvious. We have $\bar{C} \sim \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}$, $\bar{C}^* \sim \{\overline{\chi_{\gamma^{-1}}(x\gamma^{-1})}; \gamma \in \mathfrak{G}\}$, hence $\bar{C}\bar{C}^* \sim \{\rho_\gamma(x); \gamma \in \mathfrak{G}\}$ with $\rho_\gamma(x) = \sum_{\alpha \in \mathfrak{G}} \chi_\alpha(x)\overline{\chi_{\alpha\gamma^{-1}}(x\gamma^{-1})}$. Hence for $\gamma = 1$

$$\rho_1(x) \equiv \sum_{\alpha \in \mathfrak{G}} |\chi_\alpha(x)|^2. \tag{106}$$

These $\sum_{\alpha \in \mathfrak{G}}$ converge "en mesure" in the sense of (iv) in Lemma 3.7.1. But in the $\sum_{\alpha \in \mathfrak{G}}$ of (106) all terms are $\geqq 0$, hence it converges numerically (to finite or infinite values). Hence both limits agree, and (106) holds in the sense of numerical convergence.

Now $T^0(\bar{C}\bar{C}^*) = \int_S \rho_1(x)\kappa(x)\, d_\mu x$, hence (106) gives:

$$T^0(\bar{C}\bar{C}^*) \sum_{\alpha \in \mathfrak{G}} = \int_S |\chi_\alpha(x)|^2 \kappa(x)\, d_\mu x\,. \tag{107}$$

We replace $\bar{C}$ by $\bar{C}^*$, that is $\chi_\alpha(x)$ by $\overline{\chi_{\alpha^{-1}}(x\alpha^{-1})}$. Then (107) becomes, if we also replace the summation index α by α^{-1}:

$$T^0(\bar{C}^*\bar{C}) = \sum_{\alpha \in \mathfrak{G}} \int_S |\chi_\alpha(x\alpha)|^2 \kappa(x)\, d_\mu x\,. \tag{108}$$

Now by (iii) in Lemma 3.3.1 and (ii) in Lemma 3.3.3 we have always $\int_S f(x\alpha)\kappa(x)\, d_\mu x = \int_S f(x)\kappa(x)\, d_\mu x$, hence the right sides of (107) and (108) are equal. Therefore they combine to

$$T^0(\bar{C}^*\bar{C}) = T^0(\bar{C}\bar{C}^*) = \sum_{\alpha \in \mathfrak{G}} \int_S |\chi_\alpha(x)|^2 \kappa(x)\, d_\mu x\,, \tag{109}$$

which completes our proof.

LEMMA 4.2.2. (i) *If $\bar{E}$ runs over all projections $\in \mathfrak{M}$, then $T^0(\bar{E})$ is always defined, and it is a relative dimension function in the sense of* (1), p. 165, *Definition* 8.2.1. *We now assign to the relative dimension function $D_{\mathfrak{M}}(\bar{E})$ this normalization*:

$$D_{\mathfrak{M}}(\bar{E}) = T^0(\bar{E}).$$

(ii) *We have for the $\bar{E}(M)$ of* (ii) *in Definition* 3.7.1., $M \in \Gamma$:

$$D_{\mathfrak{M}}(\bar{E}(M)) = \int_M \kappa(x)\, d_\mu x = \nu(M).$$

(iii) *There exists an $M \in \Gamma$ with $\nu(M) \neq 0, \infty$.*

PROOF: We prove these statements in a somewhat changed order:
Ad (ii): We have $\bar{E}(M) = \bar{L}_{\theta_M(x)}$, where

$$\theta_M(x) \left\{ \begin{array}{l} = 1 \text{ for } x \in M \\ = 0 \text{ otherwise} \end{array} \right|.$$

Now this means (cf. also the proof of (i) in Lemma 3.7.3) $\bar{E}(M) = \bar{L}_{\theta_M(x)} \approx \{\theta_M(x)\delta_\gamma \, ; \, \gamma \in \mathfrak{G}\}$, where

$$\delta_\gamma \left\{ \begin{array}{l} = 1 \text{ for } \gamma = \mathbf{1} \\ = 0 \text{ otherwise} \end{array} \right|.$$

So by Definition 4.2.1.

$$T^0(\bar{E}(M)) = \int_S \theta_M(x)\kappa(x)\, d_\mu x = \int_M \kappa(x)\, d_\mu x = \nu(M). \tag{110}$$

As soon as we shall have proven (i) and thereby established $D\mathfrak{M}(\bar{E}(M)) = T^0(\bar{E}(M))$, (110) will go over into the formula we want to prove.

Ad (iii): Consider the $T^{(1)}, T^{(2)}, \cdots$ of (iii) in Definition 3.2.4. If all $\nu(T^{(i)}) = 0$, then summation over all $i = 1, 2, \cdots$ would give $\nu(S) = 0$, hence equivalently $\mu(S) = 0$, which is not the case. Hence there exists an $i = 1, 2, \cdots$ with $\nu(T^{(i)}) \neq 0$. Since we have also $\nu(T^{(i)}) \neq \infty$, this completes the proof.

Ad (i): Since every projection is definite, therefore for every projection $\bar{E} \in \mathfrak{M}$ $T^0(\bar{E})$ is defined (Definition 4.2.1) and $\geqq 0, \leqq +\infty$ ((i) in Lemma 4.2.1). And $T^0(\bar{E}) \neq 0, \infty$ does occur: Put $\bar{E} = \bar{E}(M)$ for an M from (iii) above, and use the already established part of (ii) above, (110): $T^0(\bar{E}(M)) = \nu(M) \neq 0, \infty$.

Therefore the character of $T^0(\bar{E})$ as a relative dimension function can be established by (**1**), p. 170, Lemma 8.3.5: It suffices to show that $T^0(\bar{E})$ fulfills the conditions (ii), (iii) in (**1**), p. 165, Definition 8.2.1.

Proof of (ii) *eod.*: We must show that $\bar{E} \sim \bar{F}(\cdots \mathfrak{M})$ implies $T^0(\bar{E}) = T^0(\bar{F})$. By (**1**), p. 151, Definition 6.1.1, and p. 141, Lemma 4.3.1, it implies $\bar{E} = \bar{U}^*\bar{U}$, $\bar{F} = \bar{U}\bar{U}^*$. Hence $T^0(\bar{E}) = T^0(\bar{F})$ by (iv) in Lemma 4.2.1.

Proof of (iii) *eod.*: We must show $T^0(\bar{E} + \bar{F}) = T^0(\bar{E}) + T^0(\bar{F})$ for orthogonal $\bar{E}, \bar{F}$. Now this is immediate by (iii) in Lemma 4.2.1.

This completes the proof.

COROLLARY: $\mathfrak{M}$ *belongs to the classes* (I) *or* (II). ($\mathfrak{G}$ *is assumed to be measurable*!)

PROOF: (ii), (iii) above together show that a projection $\bar{E} \in \mathfrak{M}$ with $D\mathfrak{M}(\bar{E}) \neq 0, \infty$ exists. So $\mathfrak{M}$ is not of class (III).—

A precise determination of the class to which $\mathfrak{M}$ belongs, when $\mathfrak{G}$ is measurable, could now be carried out. Owing to the above Corollary the possibilities are these: (I_n) ($n = 1, 2, \cdots$), (I_∞), (II_1), (II_∞). But since this discussion would coincide in all details with that one of (**1**), pp. 204–208, we omit it. The reader is referred to that work, and especially to the complete enumeration of the results in (**1**), p. 208, Theorem XII. (The rôle of μ there is played by our ν.) That theorem shows in particular that all the cases enumerated above do actually arise when S, $\mathfrak{G}$, μ are chosen suitably.

4.3. Later in this section we shall make the assumption that $\mathfrak{G}$ is not measurable.

Without knowing what the class of $\mathfrak{M}$ is, we shall nevertheless use its relative dimension function $D\mathfrak{M}(\bar{E})$. (Cf. (**1**), p. 165, Definition 8.2.1, and p. 168, Theorem VII—it does not matter now in which normalization.) We shall see ultimately, in Lemma 4.3.5, that the class is (III_∞), and hence the range of $D\mathfrak{M}(\bar{E})$ consists of 0, ∞ only—but this need not concern us now.

We begin by investigating the $D\mathfrak{M}(\bar{E}(M))$, $M \in \Gamma$ (cf. (ii) in Definition 3.7.1).

LEMMA 4.3.1. (i) *If* $M_1, M_2, \cdots \in \Gamma$ *and* $i \neq j$ *implies* $M_iM_j = 0$, *then*

$$D\mathfrak{M}(\bar{E}(M_1 + M_2 + \cdots)) = D\mathfrak{M}(\bar{E}(M_1)) + D\mathfrak{M}(\bar{E}(M_2)) + \cdots .$$

(ii) *If* $M \in \Gamma$, *then for all* $\alpha \in \mathfrak{G}$

$$D\mathfrak{M}(\bar{E}(M)) = D\mathfrak{M}(\bar{E}(M\alpha)).$$

PROOF: Ad (i): One concludes immediately from (ii) in Definition 3.7.1 that $\bar{E}(M)$ is the projection of the following closed, linear set $\mathfrak{M}(M)$:

$$\bar{E}(M) = P_{\mathfrak{M}(M)}, \qquad \mathfrak{M}(M) = (\mathfrak{F}; \mathfrak{F}(x, \alpha) = 0 \quad \text{if} \quad x \notin M). \tag{111}$$

Now it is clear that under the above assumptions on $M_1, M_2, \cdots$ $\mathfrak{M}(M_1 + M_2 + \cdots) = [\mathfrak{M}(M_1), \mathfrak{M}(M_2), \cdots]$, and the $\mathfrak{M}(M_1)$, $\mathfrak{M}(M_2)$, $\cdots$ are mutually orthogonal. Hence (**1**), p. 168, Lemma 8.3.2. gives $D\mathfrak{M}(\bar{E}(M_1 + M_2 + \cdots)) = D\mathfrak{M}(\bar{E}(M_1)) + D\mathfrak{M}(\bar{E}(M_2)) + \cdots$ as desired.

Ad (ii): $\bar{U}_\alpha$ (cf. (i) in Lemma 3.6.1) is unitary and $\in \mathfrak{I} \subset \mathfrak{R}(\mathfrak{I}) = \mathfrak{M}$, and it is clear from (111) that $\bar{U}_\alpha$ maps $\mathfrak{M}(M)$ on $\mathfrak{M}(M\alpha)$. Hence $\bar{U}_\alpha\bar{E}(M)$ is partially isometric and has the initial set $\mathfrak{M}(M)$ and the final set $\mathfrak{M}(M\alpha)$. So $\mathfrak{M}(M) \sim \mathfrak{M}(M\alpha)$ ($\cdots \mathfrak{M}$) and hence $D\mathfrak{M}(\bar{E}(M)) = D\mathfrak{M}(\bar{E}(M\alpha))$ by (**1**), p. 151, Definition 6.1.1., and p. 165, (ii) in Definition 8.2.1.

LEMMA 4.3.2. *There are only the two following alternatives*:

(a) *For no* $M \in \Gamma$ *is* $D\mathfrak{M}(\bar{E}(M)) \neq 0, \infty$.

(b) *There exists a sequence* $M_1, M_2, \cdots \in \Gamma$, $M_1 + M_2 + \cdots = S$, *such that all* $D\mathfrak{M}(\bar{E}(M_i)) \neq 0, \infty$.

PROOF: Define a subsystem K of Γ as follows: $M \in K$ if and only if there exists a sequence $M_1, M_2, \cdots \in \Gamma$ with $M_1 + M_2 + \cdots = M$, such that all $D\mathfrak{M}(\bar{E}(M_i)) \neq 0, \infty$.

Observe:

$$M^{(1)}M^{(2)}, \cdots \in K \quad \text{imply} \quad M^{(1)} + M^{(2)} + \cdots \in K. \tag{112}$$

Indeed: Form the sequence $M_1^{(i)}, M_2^{(i)}, \cdots$ for $M^{(i)}$ ($i = 1, 2, \cdots$, cf. above), then $M_j^{(i)}$, $i, j = 1, 2, \cdots$, written as a simple sequence $M_1, M_2, \cdots$ will do for $M^{(1)} + M^{(2)} + \cdots$. Next:

$$M \in K \quad \text{implies} \quad M\alpha \in K \quad \text{for every} \quad \alpha \in \mathfrak{G}. \tag{113}$$

Indeed: If $M_1, M_2, \cdots$ do for M, then $M_1\alpha, M_2\alpha, \cdots$ will do for $M\alpha$ considering (ii) in Lemma 4.3.1. And finally:

(114) If $M \in K, M \subset M'$, $\mu(M' - M) = 0$, then $M' \in K$.

Indeed: Form the sequence $M_1, M_2, \cdots$ for M, then replace M_1 by $M_1 + (M' - M)$ and leave $M_2, M_3, \cdots$ unchanged: This sequence will do for M'.

Consider now the $T^{(1)}, T^{(2)}, \cdots$ of (iii) in Definition 3.2.4 for μ. Define:

(115) $a_i =$ Least upper bound of $(\mu(M \cdot T^{(i)}); M \in K)$.

For $i, j = 1, 2, \cdots$ we can choose an $M_{i,j}$ with

(116) $M_{i,j} \in K, \quad \mu(M_{i,j} \cdot T^{(i)}) \geqq a_i - \frac{1}{j}$.

Write the $M_{i,j}$, $i, j = 1, 2, \cdots$, as a simple sequence $M^{(1)}, M^{(2)}, \cdots$. Then by (112) $\sum_{i,j=1}^{\infty} M_{i,j} = M^{(1)} + M^{(2)} + \cdots \in K$, that is

(117) $$\bar{M} = \sum_{i,j=1}^{\infty} M_{i,j} \in K.$$

We have $\mu(\bar{M} \cdot T^{(i)}) \geqq \mu(M_{i,j} \cdot T^{(i)}) \geqq a_i - \frac{1}{j}$. Since this holds for all $j = 1, 2, \cdots$, and $\mu(\bar{M} \cdot T^{(i)})$ does not depend on j, so it implies $\mu(\bar{M} \cdot T^{(i)}) \geqq a_i$. By (115), (117) also $\mu(\bar{M} \cdot T^{(i)}) \leqq a_i$, so

(118) $$\mu(\bar{M} \cdot T^{(i)}) = a_i.$$

Now consider an $M \in K$ with $M \supset \bar{M}$. Then we have $M \cdot T^{(i)} \supset \bar{M} \cdot T^{(i)}$ and both sets are $\subset T^{(i)}$, so their μ are finite. This, together with (115), (118), gives:

$$\mu((M - \bar{M}) \cdot T^{(i)}) = \mu(M \cdot T^{(i)} - \bar{M} \cdot T^{(i)})$$
$$= \mu(M \cdot T^{(i)}) - \mu(\bar{M} \cdot T^{(i)}) \leqq a_i - a_i = 0,$$

that is:

(119) $$\mu((M - \bar{M})T^{(i)}) = 0.$$

Summing over $i = 1, 2, \cdots$ and restating our hypothesis, we obtain:

(120) If $M \in K$ and $M \supset \bar{M}$, then $\mu(M - \bar{M}) = 0$.

(117), (113) give for every $\alpha \in \mathfrak{G}$ $\bar{M}\alpha \in K$. Hence by (112) $\bar{M} + \bar{M}\alpha \in K$ (put $M^{(1)} = \bar{M}, M^{(2)} = M^{(3)} = \cdots = \bar{M}\alpha$). So $M = \bar{M} + \bar{M}\alpha$ fulfills the hypotheses of (120), and therefore

(121) $$\mu((\bar{M} + \bar{M}\alpha) - \bar{M}) = 0.$$

Replacing α by α^{-1} in (121) gives $\mu((\bar{M} + \bar{M}\alpha^{-1}) - \bar{M}) = 0$, hence $\mu(((\bar{M} + \bar{M}\alpha^{-1}) - \bar{M})\alpha) = 0$, that is

$$\mu((\bar{M} + \bar{M}\alpha) - \bar{M}\alpha) = 0. \tag{122}$$

Summing (121), (122) gives:

$$\mu((\bar{M} + \bar{M}\alpha) - (\bar{M} \cdot \bar{M}\alpha)) = 0. \tag{123}$$

Since (123) holds for all $\alpha \in \mathfrak{S}$, and since $\mathfrak{S}$ is ergodic, we must either have

$$\mu(\bar{M}) = 0 \tag{124a}$$

or

$$\mu(S - \bar{M}) = 0. \tag{124b}$$

Case (124a): Consider any $M \in \Gamma$ with $D\mathfrak{M}(\bar{E}(M)) \neq 0, \infty$. Then we see, by replacing the $M_1, M_2, \cdots$ of $\bar{M}$ by $M, M_1, M_2, \cdots$, that $M + \bar{M} \in K$. $M + \bar{M}$ fulfills the hypotheses of (120), hence $\mu((M + \bar{M}) - \bar{M}) = 0$. This and (124a) give together $\mu(M + \bar{M}) = 0$, hence $\mu(M) = 0$. Now (ii) in Definition 3.7.1 shows $\bar{E}(M) = 0$ and so $D\mathfrak{M}(\bar{E}(M)) = 0$, which contradicts our assumption.

So we have never $D\mathfrak{M}(\bar{E}(M)) \neq 0, \infty$, which is the alternative (a).

Case (124b): (117), (124b), and (114) give together $S \in K$, which is the alternative (b).

Lemma 4.3.3. *If the alternative* (b) *of Lemma* 4.3.2 *holds, then* $\mathfrak{S}$ *is measurable.*

Proof: For every $M \in \Gamma$ define

$$\nu(M) = D\mathfrak{M}(\bar{E}(M)).$$

This ν is a Γ-measure, because it fulfills the conditions (i)–(iv) of Definition 3.2.4. Indeed:

Ad (i), (ii): Obvious.

Ad (iii): Immediate by (b) in Lemma 4.3.2, if we put $M_1, M_2, \cdots$ for $T^{(1)}, T^{(2)}, \cdots$.

Ad (iv): Coincides with (i) in Lemma 4.3.1.

We now prove the measurability of $\mathfrak{S}$ by showing that this ν also fulfills the conditions (i), (ii) of Definition 3.3.6. Indeed:

Ad (i): $\nu(M) = 0$ means $D\mathfrak{M}(\bar{E}(M)) = 0$, that is $\bar{E}(M) = 0$, that is, by (ii) in Definition 3.7.1, $\bar{L}_{\theta_M(x)} = 0$. But this is clearly equivalent to $\mu(M) = 0$.

Ad (ii): Coincides with (ii) in Lemma 4.3.1.

This completes the proof.—

We next derive a property of the operation $\bar{A}^{|\bar{E}_1|\bar{E}_2|\cdots}$ (cf. Definition 2.1.2).

Lemma 4.3.4. *Since* $\bar{L} \subset \mathfrak{M}$ *is an Abelian ring* (*by* (i) *in Lemma* 3.7.3), *so every sequence of projections* $\bar{E}_1, \bar{E}_2, \cdots \in \bar{\mathfrak{L}}$ *is commutative and hence for every* $\bar{A} \in \mathfrak{M}$ *of finite rank* $\bar{A}^{|\bar{E}_1|\bar{E}_2|\cdots}$ *can be formed and is* $\in \mathfrak{M}$ (*by Definition* 2.1.2).

Put $\bar{A} = \{\chi_\gamma(x); \gamma \in \mathfrak{S}\}$, $\bar{A}^{|\bar{E}_1|\bar{E}_2|\cdots} = \{\chi^*_\gamma(x); \gamma \in \mathfrak{S}\}$.

Then
$\chi_1(x) \equiv \chi_1^*(x)$, *except for an* x-*set of* μ-*measure* 0.

PROOF: All x-relations in this proof are to be understood to hold with the exception of x-sets of μ-measure 0 only.

We define

$$\bar{A}^{|\bar{E}_1|\cdots|\bar{E}_n} = \{\chi_\gamma^{(n)}(x); \gamma \in \mathfrak{G}\} \qquad \text{for } n = 1, 2, \cdots, \tag{125}$$

and proceed in several successive steps:

Proof of $\chi_1(x) \equiv \chi_1^{(1)}(x)$: We have

$$\bar{A}^{|\bar{E}_1} = \bar{E}_1\bar{A}\bar{E}_1 + (I - \bar{E}_1)\bar{A}(I - \bar{E}). \tag{126}$$

Now $\bar{E}_1 \in \overline{\mathfrak{L}}$, so by (ii) in Lemma 3.7.3, $\bar{E}_1 = \bar{E}(M_1)$ for an $M_1 \in \Gamma$. That is, $\bar{E}_1 = \bar{L}_{\theta_{M_1}(x)}$,

$$\theta_{M_1}(x) \left\{ \begin{array}{l} = 1 \text{ for } x \in M_1 \\ = 0 \text{ otherwise} \end{array} \right|.$$

This means, as we know, $\bar{E}_1 = \{\theta_{M_1}(x)\delta_\gamma ; \gamma \in \mathfrak{G}\}$,

$$\delta_\gamma \left\{ \begin{array}{l} = 1 \text{ for } \gamma = \mathbf{1} \\ = 0 \text{ otherwise} \end{array} \right|.$$

This and $\bar{A} = \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}$ give, owing to (126):

$$\bar{A}^{|\bar{E}_1} = \{[\theta_{M_1}(x)\theta_{M_1}(x\gamma^{-1}) + (1 - \theta_{M_1}(x))\,(1 - \theta_{M_1}(x\gamma^{-1})]\chi_\gamma(x); \gamma \in \mathfrak{G}\}.$$

Hence by (125):

$$\chi_\gamma^{(1)}(x) \equiv [\theta_{M_1}(x)\theta_{M_1}(x\gamma^{-1}) + (1 - \theta_{M_1}(x))\,(1 - \theta_{M_1}(x\gamma^{-1}))]\chi_\gamma(x). \tag{127}$$

Put $\gamma = \mathbf{1}$ in (127); then, since $\theta_M(x) = 0, 1$, it goes over into $\chi_1^{(1)}(x) \equiv \chi_1(x)$.

Proof of $\chi_1^{(n)}(x) \equiv \chi_1^{(n+1)}(x)$ (*for* $n = 1, 2, \cdots$): We proved $\chi_1(x) \equiv \chi_1^{(1)}(x)$. Replace $\bar{A}$, $\bar{E}_1$ by $\bar{A}^{|\bar{E}_1|\cdots|\bar{E}_n}$, $\bar{E}_{n+1}$, then this becomes $\chi_1^{(n)}(x) \equiv \chi_1^{(n+1)}(x)$.

Combining the two above results gives:

$$\chi_1(x) \equiv \chi_1^{(1)}(x) \equiv \chi_1^{(2)}(x) \equiv \cdots. \tag{128}$$

Proof of $\chi_1(x) \equiv \chi_1^*(x)$: We have (by Definition 2.1.2)

$$\lim_{n\to\infty} \operatorname{str} \bar{A}^{|\bar{E}_1|\cdots|\bar{E}_n} = \bar{A}^{|\bar{E}_1|\bar{E}_2|\cdots}.$$

That is: For every $\mathcal{F} \in \mathfrak{H}_s^{\mathfrak{G}}$ $\lim_{n\to\infty} ||\bar{A}^{|\bar{E}_1|\cdots|\bar{E}_n}\mathcal{F} - \bar{A}^{|\bar{E}_1|\bar{E}_2|\cdots}\mathcal{F}||_s^{\mathfrak{G}} = 0$, and hence for all $\mathcal{F}, \mathcal{H} \in \mathfrak{H}_s^{\mathfrak{G}}$

$$\lim_{n\to\infty} ((\bar{A}^{|\bar{E}_1|\cdots|\bar{E}_n}\mathcal{F}, \mathcal{H})_s^{\mathfrak{G}} - (\bar{A}^{|\bar{E}_1|\bar{E}_2|\cdots}\mathcal{F}, \mathcal{H})) = 0. \tag{129}$$

Choose any $f, g \in \mathfrak{H}_S$, and define

$$\mathcal{F}(x, \alpha) \begin{cases} = f(x) \text{ for } \alpha = 1 \\ = 0 \quad \text{otherwise} \end{cases},$$

$$\mathcal{K}(x, \alpha) \begin{cases} = g(x) \text{ for } \alpha = 1 \\ = 0 \quad \text{otherwise} \end{cases},$$

then $\mathcal{F}, \mathcal{K} \in \mathfrak{H}_S^{\mathfrak{S}}$, and

$$(\bar{A}^{|E_1|\cdots|E_n}\mathcal{F}, \mathcal{K})_S^{\mathfrak{S}} = \int_S \chi_1^{(n)}(x) f(x)\overline{g(x)}\, d_\mu x,$$

$$(\bar{A}^{|E_1|E_2|\cdots}\mathcal{F}, \mathcal{K})_S^{\mathfrak{S}} = \int_S \chi_1^{*}(x) f(x)\overline{g(x)}\, d_\mu x.$$

Hence (129) becomes

$$\lim_{n\to\infty} \int_S (\chi_1^{(n)}(x) - \chi_1^{*}(x)) f(x)\overline{g(x)}\, d_\mu x = 0,$$

or, considering (128),

$$\int_S (\chi_1(x) - \chi_1^{*}(x)) f(x)\overline{g(x)}\, d_\mu x = 0. \tag{130}$$

If $M \in \Gamma$, $\mu(M)$ finite, we can put

$$f(x) \begin{cases} = 1 & \text{for} \quad x \in M \\ = 0 & \text{otherwise} \end{cases},$$

$$g(x) \begin{cases} = \operatorname{sgn}(\chi_1(x) - \chi_1^{*}(x)) & \text{for} \quad x \in M \\ = 0 & \text{otherwise} \end{cases},$$

(cf. footnote[12]), then $f, g \in \mathfrak{H}_S$, and (130) becomes

$$\int_M |\chi_1(x) - \chi_1^{*}(x)|\, d_\mu x = 0. \tag{131}$$

Thus $\chi_1(x) \equiv \chi_1^{*}(x)$ for all $x \in M$ (exceptions as above!). Now with the $T^{(1)}$, $T^{(2)}, \cdots$ of (iii) in Definition 3.2.4 we may put $M = T^{(i)}$, and sum over $i = 1, 2, \cdots$; this gives $\chi_1(x) \equiv \chi_1^{*}(x)$ unrestrictedly.

Thus the proof is completed.—

We are now in the position to take the decisive step.

LEMMA 4.3.5. *If $\mathfrak{S}$ is not measurable, then $\mathfrak{M}$ is of class* (III_∞).

PROOF: Assume that $\mathfrak{S}$ is not measurable. We shall show that $\mathfrak{M}$ is of class (III_∞) by applying Lemma 2.2.2. For the $\mathfrak{A}$ mentioned there we shall choose the $\overline{\mathfrak{L}}$ of (i) in Definition 3.7.1. So we must only prove the hypotheses of Lemma 2.2.2 for $\mathfrak{A} = \overline{\mathfrak{L}}$.

$\overline{\mathfrak{L}}$ is an Abelian ring $\subset \mathfrak{M}$: By (i) in Lemma 3.7.3.

Then
$\chi_1(x) \equiv \chi_1^*(x)$, *except for an* x*-set of* μ*-measure* 0.

PROOF: All x-relations in this proof are to be understood to hold with the exception of x-sets of μ-measure 0 only.

We define

$$\bar{A}^{|\bar{E}_1|\cdots|\bar{E}_n} = \{\chi_\gamma^{(n)}(x); \gamma \in \mathfrak{G}\} \qquad \text{for } n = 1, 2, \cdots, \tag{125}$$

and proceed in several successive steps:

Proof of $\chi_1(x) \equiv \chi_1^{(1)}(x)$: We have

$$\bar{A}^{|\bar{E}_1} = \bar{E}_1\bar{A}\bar{E}_1 + (I - \bar{E}_1)\bar{A}(I - \bar{E}). \tag{126}$$

Now $\bar{E}_1 \in \overline{\mathfrak{L}}$, so by (ii) in Lemma 3.7.3, $\bar{E}_1 = \bar{E}(M_1)$ for an $M_1 \in \Gamma$. That is, $\bar{E}_1 = \overline{L}_{\theta_{M_1}(x)}$,

$$\theta_{M_1}(x) \begin{cases} = 1 \text{ for } x \in M_1 \\ = 0 \text{ otherwise} \end{cases}.$$

This means, as we know, $\bar{E}_1 \approx \{\theta_{M_1}(x)\delta_\gamma ; \gamma \in \mathfrak{G}\}$,

$$\delta_\gamma \begin{cases} = 1 \text{ for } \gamma = \mathbf{1} \\ = 0 \text{ otherwise} \end{cases}.$$

This and $\bar{A} \approx \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}$ give, owing to (126):

$$\bar{A}^{|\bar{E}_1} \approx \{[\theta_{M_1}(x)\theta_{M_1}(x\gamma^{-1}) + (1 - \theta_{M_1}(x))\,(1 - \theta_{M_1}(x\gamma^{-1})]\chi_\gamma(x); \gamma \in \mathfrak{G}\}.$$

Hence by (125):

$$\chi_\gamma^{(1)}(x) \equiv [\theta_{M_1}(x)\theta_{M_1}(x\gamma^{-1}) + (1 - \theta_{M_1}(x))\,(1 - \theta_{M_1}(x\gamma^{-1}))]\chi_\gamma(x). \tag{127}$$

Put $\gamma = \mathbf{1}$ in (127); then, since $\theta_M(x) = 0, 1$, it goes over into $\chi_1^{(1)}(x) \equiv \chi_1(x)$.

Proof of $\chi_1^{(n)}(x) \equiv \chi_1^{(n+1)}(x)$ *(for* $n = 1, 2, \cdots$*)*: We proved $\chi_1(x) \equiv \chi_1^{(1)}(x)$. Replace $\bar{A}$, $\bar{E}_1$ by $\bar{A}^{|\bar{E}_1|\cdots|\bar{E}_n}$, $\bar{E}_{n+1}$, then this becomes $\chi_1^{(n)}(x) \equiv \chi_1^{(n+1)}(x)$.

Combining the two above results gives:

$$\chi_1(x) \equiv \chi_1^{(1)}(x) \equiv \chi_1^{(2)}(x) \equiv \cdots. \tag{128}$$

Proof of $\chi_1(x) \equiv \chi_1^*(x)$: We have (by Definition 2.1.2)

$$\lim_{n\to\infty} \operatorname{str} \bar{A}^{|\bar{E}_1|\cdots|\bar{E}_n} = \bar{A}^{|\bar{E}_1|\bar{E}_2|\cdots}.$$

That is: For every $\mathcal{F} \in \mathfrak{H}_s^{\mathfrak{G}}$ $\lim_{n\to\infty} \| \bar{A}^{|\bar{E}_1|\cdots|\bar{E}_n}\mathcal{F} - \bar{A}^{|\bar{E}_1|\bar{E}_2|\cdots}\mathcal{F}\|_s^{\mathfrak{G}} = 0$, and hence for all $\mathcal{F}, \mathcal{H} \in \mathfrak{H}_s^{\mathfrak{G}}$

$$\lim_{n\to\infty} ((\bar{A}^{|\bar{E}_1|\cdots|\bar{E}_n}\mathcal{F}, \mathcal{H})_s^{\mathfrak{G}} - (\bar{A}^{|\bar{E}_1|\bar{E}_2|\cdots}\mathcal{F}, \mathcal{H})) = 0. \tag{129}$$

Choose any $f, g \in \mathfrak{H}_S$, and define

$$\mathcal{F}(x, \alpha) \begin{cases} = f(x) \text{ for } \alpha = 1 \\ = 0 \quad \text{otherwise} \end{cases},$$

$$\mathcal{H}(x, \alpha) \begin{cases} = g(x) \text{ for } \alpha = 1 \\ = 0 \quad \text{otherwise} \end{cases},$$

then $\mathcal{F}, \mathcal{H} \in \mathfrak{H}_S^{\mathfrak{S}}$, and

$$(\bar{A}^{|E_1|\cdots|E_n}\mathcal{F}, \mathcal{H})_S^{\mathfrak{S}} = \int_S \chi_1^{(n)}(x) f(x)\overline{g(x)}\, d_\mu x,$$

$$(\bar{A}^{|E_1|E_2|\cdots}\mathcal{F}, \mathcal{H})_S^{\mathfrak{S}} = \int_S \chi_1^{*}(x) f(x)\overline{g(x)}\, d_\mu x.$$

Hence (129) becomes

$$\lim_{n\to\infty} \int_S (\chi_1^{(n)}(x) - \chi_1^{*}(x)) f(x)\overline{g(x)}\, d_\mu x = 0,$$

or, considering (128),

$$\int_S (\chi_1(x) - \chi_1^{*}(x)) f(x)\overline{g(x)}\, d_\mu x = 0. \tag{130}$$

If $M \in \Gamma$, $\mu(M)$ finite, we can put

$$f(x) \begin{cases} = 1 & \text{for} \quad x \in M \\ = 0 & \text{otherwise} \end{cases},$$

$$g(x) \begin{cases} = \operatorname{sgn}(\chi_1(x) - \chi_1^{*}(x)) & \text{for} \quad x \in M \\ = 0 & \text{otherwise} \end{cases}$$

(cf. footnote[12]), then $f, g \in \mathfrak{H}_S$, and (130) becomes

$$\int_M |\chi_1(x) - \chi_1^{*}(x)|\, d_\mu x = 0. \tag{131}$$

Thus $\chi_1(x) \equiv \chi_1^{*}(x)$ for all $x \in M$ (exceptions as above!). Now with the $T^{(1)}$, $T^{(2)}, \cdots$ of (iii) in Definition 3.2.4 we may put $M = T^{(i)}$, and sum over $i = 1, 2, \cdots$; this gives $\chi_1(x) \equiv \chi_1^{*}(x)$ unrestrictedly.

Thus the proof is completed.—

We are now in the position to take the decisive step.

LEMMA 4.3.5. *If $\mathfrak{S}$ is not measurable, then $\mathfrak{M}$ is of class* (III_∞).

PROOF: Assume that $\mathfrak{S}$ is not measurable. We shall show that $\mathfrak{M}$ is of class (III_∞) by applying Lemma 2.2.2. For the $\mathfrak{A}$ mentioned there we shall choose the $\overline{\mathfrak{L}}$ of (i) in Definition 3.7.1. So we must only prove the hypotheses of Lemma 2.2.2 for $\mathfrak{A} = \overline{\mathfrak{L}}$.

$\overline{\mathfrak{L}}$ is an Abelian ring $\subset \mathfrak{M}$: By (i) in Lemma 3.7.3.

Ad (i): The ring $\mathfrak{M}\cdot\overline{\mathfrak{L}'}(\subset\mathfrak{M})$ is purely infinite: By Lemmata 4.3.2, 4.3.3, since $\mathfrak{G}$ is not measurable, the alternative (a) of Lemma 4.3.2 holds. But this means, considering (ii) in Lemma 3.7.3 and Definition 2.2.1, that $\overline{\mathfrak{L}}$ is purely infinite.

Now by (i) in Lemma 3.7.3, $\overline{\mathfrak{L}} = \mathfrak{M}\cdot\overline{\mathfrak{L}'}$, hence we have shown that $\mathfrak{M}\cdot\overline{\mathfrak{L}'}$ is purely infinite.

Ad (ii): For every finite projection $\bar{F}(\epsilon\mathfrak{M})$ $\bar{F}^{|\bar{E}_1|\bar{E}_2|\cdots} = 0$ implies $\bar{F} = 0$: Assume $\bar{F}$ as described. Put $\bar{F} \approx \{\chi_\gamma(x);\ \gamma\ \epsilon\ \mathfrak{G}\}$, $\bar{F}^{|\bar{E}_1|\bar{E}_2|\cdots} \approx \{\chi^*_\gamma(x);\ \gamma\ \epsilon\ \mathfrak{G}\}$. Then $\bar{F}^{|\bar{E}_1|\bar{E}_2|\cdots} = 0$ implies $\chi^*_1(x) \equiv 0$. (All x-relations are to be understood to hold with the exception of x-sets of μ-measure 0 only.) By Lemma 4.3.4, $\chi_1(x) \equiv \chi^*_1(x)$, hence $\chi_1(x) \equiv 0$. Now $\bar{F}$ being a projection, is definite. Hence $\chi_1(x) \equiv 0$ implies, by (iii) in Lemma 3.7.2, that $\bar{F} = 0$.

Thus the proof is completed.—

We may sum up our results as follows:

THEOREM IX. *Let S, Γ, μ be as in Theorem* VIII, *and form $\mathfrak{H}_S$, $\mathfrak{H}_S^{\mathfrak{G}}$ and $\mathfrak{M}$, $\mathfrak{M}'$ as described there. Then we have:*
If $\mathfrak{G}$ is measurable (cf. Definition 3.3.6), *then the factors $\mathfrak{M}$, $\mathfrak{M}'$ belong to the classes* (I) *or* (II). *The precise determination of their classes is given in* (**1**), *p.* 204, *Lemma* 13.1.1, *and p.* 206, *Lemma* 13.2.1. *(Cf. the end of our* §4.2.) *If $\mathfrak{G}$ is not measurable, then the factors $\mathfrak{M}$,$\mathfrak{M}'$ belong to the class* (III) *(that is,* (III_∞)).

PROOF: It suffices to consider $\mathfrak{M}$ only: The automorphism $\mathcal{F} \rightleftarrows \overline{W}\mathcal{F}$ of $\mathfrak{H}_S^{\mathfrak{G}}$ (cf. Theorem VIII) extends our results from $\mathfrak{M}$ to $\mathfrak{M}'$.

Our statements concerning a measurable $\mathfrak{G}$ result from the Corollary to Lemma 4.2.2, and the discussion which follows, at the end of §4.2. Those concerning a non-measurable $\mathfrak{G}$ result from Lemma 4.3.5.

4.4. Examples of S, Γ, μ and $\mathfrak{G}$ as required by Theorems VIII, IX for any one of the classes (I)–(II)—that is, of free, ergodic, and measurable groups $\mathfrak{G}$—were given in (**1**), pp. 206–208, Lemma 13.2.1, and the specific examples (α) – (γ). So there is only one more thing left for us to do: To give examples of S, Γ, μ and $\mathfrak{G}$ as required by Theorems VIII, IX for the class (III)—that is: of free, ergodic, and non-measurable goups $\mathfrak{G}$.

A certain class of examples was mentioned in the Introduction, §4, but we shall not investigate it here. Our examples will be based on the following lemma:

LEMMA 4.4.1. *Let S, Γ, μ and $\mathfrak{G}$ be as described in Theorem* VIII. *Denote the set of those $\alpha\ \epsilon\ \mathfrak{G}$, for which always $\mu(M) = \mu(M\alpha)$, by $\mathfrak{G}_0$. $\mathfrak{G}_0$ is a subgroup of $\mathfrak{G}$ and a free S, Γ, μ-group along with $\mathfrak{G}$. It is also measurable, with μ for ν. (Cf. Definitions* 3.3.3, 3.3.4, *and* 3.3.6.) *Now we assume:*
(i) *$\mathfrak{G}_0$ is ergodic. (Cf. Definition* 3.3.5.)
(ii) $\mathfrak{G}_0 \neq \mathfrak{G}$.
Then $\mathfrak{G}$ is not measurable.

PROOF: The statements about $\mathfrak{G}_0$ which precede (i), (ii) are obvious on the basis of the definitions referred to.

Assume now (i), (ii), and, *per absurdum*, the measurability of $\mathfrak{G}$. Consider the ν of Definition 3.3.6 for $\mathfrak{G}$. This ν will also do for $\mathfrak{G}_0$. So μ, ν both do for $\mathfrak{G}_0$. Now $\mathfrak{G}_0$ is ergodic by (i), hence Lemma 3.3.4 applies for $\mathfrak{G}_0$ in place of $\mathfrak{G}$, and gives $\mu(M) = a\nu(M)$ for a constant factor $a > 0$. Since we have for every $\alpha \in \mathfrak{G}$ $\nu(M) = \nu(M\alpha)$ by the definition of ν, so this gives $\mu(M) = \mu(M\alpha)$, that is $\alpha \in \mathfrak{G}_0$. Hence $\mathfrak{G}_0 = \mathfrak{G}$, which contradicts (ii).

This completes the proof.—

We can now give specific examples.

EXAMPLE (α). Let S be the set of all real numbers, Γ the system of all Borel sets in S (in the usual sense), and μ the common Lebesgue measure (for Borel sets). Consider the following one-to-one mappings of S on itself:

$$\alpha_1(\rho, \sigma) : x \rightleftarrows \rho x + \sigma, \qquad \rho > 0,\ \rho, \sigma \text{ rational.} \tag{132}$$

These $\alpha_1(\rho, \sigma)$ form an obviously enumerably infinite group. Let $\mathfrak{G}$ be this group. One verifies immediately that $\mathfrak{G}$ is an S, Γ, μ-group, and also that it is free.

Clearly $\mu(M\alpha) = \rho\mu(M)$ for $\alpha = \alpha_1(\rho, \sigma)$, hence the set $\mathfrak{G}_0$ of all $\alpha \in \mathfrak{G}$ with $\mu(M\alpha) = \mu(M)$ consists of all $\alpha = \alpha_1$, (ρ, σ) with $\rho = 1$. So $\mathfrak{G}_0 \neq \mathfrak{G}$.

$\mathfrak{G}_0$ is also an S, Γ, μ-group, and ergodic by (1), p. 206, Lemma 13.2.1, and p. 208, example (β). Hence $\mathfrak{G}$ is a fortiori ergodic.

So S, Γ, μ and $\mathfrak{G}$ fulfill the requirements of Theorems VIII, IX, and $\mathfrak{G}_0$ fulfills the conditions of Lemma 4.4.1. Hence $\mathfrak{G}$ is not measurable.

Observe that in this case $\mu(S) = \infty$.

EXAMPLE (β): Let S be the set of all complex numbers z with $|z| = 1$, Γ the system of all Borel sets in S (in the usual sense). Let μ be the common Lebesgue measure to be applied after S has been mapped by the mapping $z = e^{2\pi i\varphi}$ on the set S, of all real numbers $\varphi \geqq 0, < 1$ which can be (and in what follows will be) looked at as the set of all real numbers (mod 1).

Consider the following one-to-one mappings of S on itself:

$$\alpha_2(\theta, \mu) : z \rightleftarrows \theta \frac{z + u}{1 + \bar{u}z}, \qquad |\theta| = 1, |u| < 1. \tag{133}$$

These $\alpha_2(\theta, u)$ form an obviously continuous group. Denote this group by $\bar{\mathfrak{G}}$.[13] Consider the following elements of $\bar{\mathfrak{G}}$: The $\alpha = \alpha_2(\theta, u)$ with $\theta = e^{2\pi i\rho}$, ρ rational, $u = 0$, and that one with $\theta = 1$, $u = \frac{1}{2}$. They form an enumerably infinite set, and consequently they generate an enumerably infinite subgroup of $\bar{\mathfrak{G}}$. Let $\mathfrak{G}$ be this subgroup.

In combining two $\alpha = \alpha_2(\theta, u)$ of (133), their θ simply multiply. Hence the θ of all $\alpha = \alpha_2(\theta, u) \in \mathfrak{G}$ are the products of the θ of those $\alpha = \alpha_2(\theta, u)$ with which we started, that is: the $\theta = e^{2\pi i\rho}$, ρ rational. Therefore we have:

$$\alpha = \alpha_2(\theta, 0) \in \mathfrak{G} \text{ if and only if } \theta = e^{2\pi i\rho},\ \rho \text{ rational.} \tag{134}$$

One verifies immediately that $\mathfrak{G}$ is an S, Γ, μ-group, and also that it is free.

[13] $\bar{\mathfrak{G}}$ plays a well-known and important rôle in the theory of conformal mappings.

Let us now determine $\mathfrak{G}_0$, $\alpha \in \mathfrak{G}_0$ means that $\mu(M\alpha) = \mu(M)$. Since $\alpha = \alpha_2(\theta, u)$ maps arcs on arcs, this requires that it conserve their lengths, hence, that it be a rotation or a reflection. That is:

$$\alpha\colon z \rightleftarrows \theta z, \qquad \text{or} \qquad z \rightleftarrows \theta \frac{1}{z}, \qquad |\theta| = 1.$$

(133) excludes the second alternative, and the first one means $u = 0$ for $\alpha = \alpha_2(\theta, u)$. So $\mathfrak{G}_0$ is characterized by $u = 0$, and consequently by (134). So $\mathfrak{G} \neq \mathfrak{G}_0$, since $\alpha = \alpha_2(1, \frac{1}{2})$ is $\in \mathfrak{G}$ but not $\in \mathfrak{G}_0$.

Now map S by $z = e^{2\pi i\varphi}$ on the φ-set S, (cf. above). Then the mappings of $\mathfrak{G}_0$ go over from $z \rightleftarrows e^{2\pi i\rho} z$ (in (134)) into $\varphi \rightleftarrows \varphi + \rho$, ρ rational. This proves that $\mathfrak{G}_0$ is ergodic (in S_1, and hence also in S), by (**1**), p. 206, Lemma 13.2.1, and p. 208, example (β). Hence $\mathfrak{G}$ is a fortiori ergodic.

So S, Γ, μ and $\mathfrak{G}$ fulfill all requirements of Theorems VIII, IX, and $\mathfrak{G}_0$ fulfills the conditions of Lemma 4.4.1. Hence $\mathfrak{G}$ is not measurable.

Observe that in this case $\mu(S) = 1$.—

Besides establishing the existence of factors $\mathfrak{M}$ of class (III), the above examples can also be used to determine all existing types of factorizations $\mathfrak{M}_1$, $\mathfrak{M}_2$. To this end we need the lemma which follows.

LEMMA 4.4.2. *Consider the* S, Γ, μ, $\mathfrak{G}$, $\mathfrak{G}_0$ *and* $\mathfrak{M}$, $\mathfrak{M}'$ *of Lemma* 4.4.1. *Every* $\bar{A} \in \mathfrak{M}'$ *is* $\bar{A} \approx \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}$; *consider those for which* $\chi_\gamma(x) \equiv 0$ *whenever* $\gamma \notin \mathfrak{G}_0$. *Denote their set by* $\mathfrak{N}$. *Then we have*:

(i) $\mathfrak{N}$ *is a ring.*

(ii) $\mathfrak{M}' \cdot \mathfrak{N}' = (aI)$ (*a runs over all complex numbers*).

(iii) $\mathfrak{M}$, $\mathfrak{N}$ *is a factorization, but* $\mathfrak{M}$, $\mathfrak{N}$ *are not coupled factors.* (Cf. (**1**), p. 138, *Definition* 3.1.3.)

(iv) *The class of the factor* $\mathfrak{N}$ *is* (II_1) *or* (II_∞) *according to whether* $\mu(S)$ *is finite or not.*

PROOF: Ad (i), (ii): Literally the same as the proof of (**1**), pp. 208–209, Lemma 13.4.1.

Ad (iii): Literally the same as the proof of (**1**), p. 209, Lemma 13.4.2.

Ad (iv): For $\bar{A} \in \mathfrak{N}$, $\bar{A}$ definite, we have also $\bar{A} \in \mathfrak{M}$, $\bar{A} \approx \{\chi_\gamma(x); \gamma \in \mathfrak{G}\}$ (but $\chi_\gamma(x) \equiv 0$ for $\gamma \notin \mathfrak{G}_0$!), and we define again $T_0(\bar{A}) = \int_S \chi_1(x)\, d_\mu x$. Then the considerations of Lemmas 4.2.1., 4.2.2, apply literally in $\mathfrak{N}$ (because for $\alpha \in \mathfrak{G}_0$ $\mu(M) = \mu(M\alpha)$, and for all other $\alpha \in \mathfrak{G}$ we have $\chi_\alpha(x) \equiv 0$). Hence $D_{\mathfrak{N}}(\bar{E}) = T_0(\bar{E})$ for all projections $\bar{E} \in \mathfrak{N}$. Hence we see, as in the Corollary after Lemma 4.2.2, that $\mathfrak{N}$ belongs to the classes (I) or (II).

If $\mathfrak{N}$ were of class (I), then $\mathfrak{N}'$ would be a normal factor by (**1**), p. 184, Lemma 11.2.2, and hence $\mathfrak{M}$, $\mathfrak{N}$ would be a coupled factorization by (**1**), p. 183, Lemma 11.1.2. This contradicts (iii) above, hence $\mathfrak{N}$ is of class (II).

Now $D_{\mathfrak{N}}(I) = T_0(I) = \int_S 1 \cdot d_\mu x = \mu(S)$, hence $\mathfrak{N}$ is of class (II_1) or (II_∞), according to whether $\mu(S)$ is finite or not.—

Thus Lemma 4.4.2 shows, if we avail ourselves of the above Examples (α), (β), that non-coupled factorizations of the classes (II_∞), (III_∞) and (II_1), (III_∞) exist. This shows that no analogue of (1), p. 182, Theorem X (which holds for coupled factorizations) holds for non-coupled factorizations. Indeed, the only restriction on non-coupled factorizations is that no factor can be of class (I) (by (1), pp. 183–184, Lemmas 11.1.2, 11.2.2, cf. the above proof of (iv) in Lemma 4.4.2). We could also show the existence of non-coupled factorization of class (III_∞), (III_∞), but we shall not discuss this now.

INSTITUTE FOR ADVANCED STUDY.

ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

ON THE UNIFORMLY-DENSE DISTRIBUTION OF CERTAIN SEQUENCES OF POINTS

BY PAUL ERDÖS AND PAUL TURÁN

(Received December 15, 1938)

Let

$$\mathfrak{M}_1 \equiv \left\{ \begin{matrix} \varphi_1^{(1)} \\ \varphi_1^{(2)}, \varphi_2^{(2)} \\ \vdots \\ \varphi_1^{(n)}, \cdots \varphi_n^{(n)} \\ \vdots \qquad \ddots \end{matrix} \right\} \tag{1}$$

be a sequence of *numbers*, where for all n

$$0 \leqq \varphi_1^{(n)} < \varphi_2^{(n)} < \cdots < \varphi_n^{(n)} \leqq \pi.$$

Weyl[1] calls such a sequence uniformly-dense in $[0, 2\pi]$, if for every subinterval $[\alpha, \beta]$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\alpha \leqq \varphi_\nu^{(n)} \leqq \beta} 1 = \frac{\beta - \alpha}{\pi}. \tag{2}$$

Weyl proved that the sequence $\mathfrak{M}_1$ is uniformly dense in $[0, 2\pi]$ if and only if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} e^{ki\varphi_\nu^{(n)}} = 0 \tag{3}$$

for every integer k.

Suppose we are given on the unit circle of the complex z-plane a sequence of *points*

$$\mathfrak{M}_2 \equiv \left\{ \begin{matrix} z_1^{(1)} \\ z_1^{(2)}, z_2^{(2)} \\ \vdots \\ z_1^{(n)}, z_2^{(n)}, \cdots z_n^{(n)} \\ \vdots \quad \vdots \qquad \ddots \end{matrix} \right\} \equiv \left\{ \begin{matrix} e^{i\varphi_1^{(1)}} \\ e^{i\varphi_1^{(2)}}, e^{i\varphi_2^{(2)}} \\ \vdots \\ e^{i\varphi_1^{(n)}}, \cdots e^{i\varphi_n^{(n)}} \\ \vdots \qquad \ddots \end{matrix} \right\}$$

with

$$0 \leqq \varphi_1^{(n)} < \varphi_2^{(n)} < \cdots < \varphi_n^{(n)} \leqq 2\pi$$

[1] H. Weyl: *Über die Gleichverteilung von Zahlen mod Eins*, Math. Ann., 1916, Bd. 77, pp. 313–352.

for all n. The sequence is called uniformly dense on the unit circle, if for any arc of the length l

$$\lim_{n\to\infty}\frac{1}{n}\sum_{\substack{\nu\\ s_\nu^{(n)}<l}} 1=\frac{l}{2\pi}, \tag{4}$$

holds. It is satisfied if and only if the sequence of *numbers*

$$\left\{\begin{array}{c}\varphi_1^{(1)}\\ \vdots\\ \vdots\\ \varphi_1^{(n)}\ \cdots\ \varphi_n^{(n)}\\ \vdots\qquad\ddots\end{array}\right\}$$

is, in the sense of (2), uniformly dense in $[0, 2\pi]$.

Let a closed Jordan-curve l of the complex ζ-plane be given. The sequence of points

$$\mathfrak{M}_3\equiv\left\{\begin{array}{c}\zeta_1^{(1)}\\ \vdots\\ \vdots\\ \zeta_1^{(n)},\ \zeta_2^{(n)},\ \cdots\ \zeta_n^{(n)}\\ \vdots\quad\vdots\qquad\ddots\end{array}\right\}$$

lying on l is called uniformly dense if, mapping the exterior of l schlicht-conformally and the periphery continuously on the closed exterior of the unit circle of the z-plane, we obtain on the circumference of the unit circle a sequence of points

$$\left\{\begin{array}{c}z_1^{(1)}\\ \vdots\\ \vdots\\ z_1^{(n)},\ z_2^{(n)},\ \cdots\ z_n^{(n)}\\ \vdots\quad\vdots\qquad\ddots\end{array}\right\} \tag{5}$$

uniformly dense in the sense of (4).[2]

We have to explain the case, when l degenerates into an open arc on the ζ-plane. In this case, in the same say as in (5), we obtain two $z_\nu^{(n)}$ belonging to one $\zeta_\nu^{(n)}$. Let l be e.g. the interval $[-1, +1]$; in this case the mapping function being $\zeta=\frac{1}{2}\left(z+\frac{1}{z}\right)$, the connection between the $\zeta_\nu^{(n)}$ and $z_\nu^{(n)}$, i.e. between $\zeta_\nu^{(n)}$ and $\varphi_\nu^{(n)}$, respectively, is

$$\zeta_\nu^{(n)}=\tfrac{1}{2}(e^{i\varphi_\nu^{(n)}}+e^{-i\varphi_\nu^{(n)}})=\cos\varphi_\nu^{(n)}$$

$$\nu=1, 2, \cdots n, \qquad n=1, 2, \cdots, \qquad 0\leqq\varphi_\nu^{(n)}<2\pi.$$

[2] The uniform-dense distribution of points lying on ρ, but with potential theoretic characterization, occurs at first in the investigations of Hilbert; see D. Hilbert: *Über die Entwickelung einer beliebigen analytischen Function einer Variabeln* ⋯ , Nachrichten von der Königlichen Gesellschaft Göttingen, 1897, pp. 63–70. The above formalization, which is equivalent to Hilbert's, is due to L. Fejèr, see L. Fejèr: *Interpolation und konforme Abbildung*, Nachrichten van der Kön. Ges. Göttingen, 1918, pp. 319–331.

For fixed ν and n, $\varphi_\nu^{(n)}$ has two values, which lie symmetrically with respect to $\varphi = \pi$; so *we call the sequence of points*

$$\left\{\begin{matrix} \zeta_1^{(1)} \\ \vdots \\ \zeta_1^{(n)}, \cdots \zeta_n^{(n)} \\ \vdots \qquad \ddots \end{matrix}\right\} \tag{6}$$

lying on $(-1, +1]$, *with*

$$1 \geqq \zeta_1^{(n)} > \zeta_2^{(n)} > \cdots > \zeta_n^{(n)} \geqq -1, \qquad n = 1, 2, \cdots,$$

uniformly dense, if for the sequence of numbers

$$\left\{\begin{matrix} \varphi_1^{(1)} \\ \vdots \\ \varphi_1^{(n)}, \varphi_2^{(n)}, \cdots \varphi_n^{(n)} \\ \vdots \quad \vdots \qquad \ddots \end{matrix}\right\}$$

defined by[3]

$$\zeta_\nu^{(n)} = \cos \varphi_\nu^{(n)}, \; 0 \leqq \varphi_\nu^{(n)} \leqq \pi, \; \nu = 1, 2, \cdots n, \; n = 1, 2, \cdots,$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\alpha \leqq \varphi_\nu^{(n)} \leqq \beta} 1 = \frac{\beta - \alpha}{\pi},$$

holds, where $[\alpha, \beta]$ *means an arbitrary subinterval of* $[0, \pi]$.

To this definition, appearing artificial at the first moment, we may give geometric sense in the following way: Let us draw upon $[-1, +1]$ a semicircle and project the point $x_\nu^{(n)}$ upon the circle and obtain $A_\nu^{(n)}$. Then $\varphi_\nu^{(n)}$ means clearly the angle between the positive real axis and $0A_\nu^{(n)}$. So the above definition means that a sequence of points in $[-1, +1]$ is here called uniformly dense if, *projected upon the unit circle, the projections* are uniformly distributed. According to this definition the most uniform distribution is the case $\varphi_\nu^{(n)} = \nu\pi/(n+1)$, $(\nu = 1, 2, \cdots n, \; n = 1, 2, \cdots)$, attained when the sequence of points in $[-1, +1]$ is $\zeta_\nu^{(n)} = \cos \nu\pi/(n+1)$ $(\nu = 1, 2, \cdots, n, \; n = 1, 2, \cdots)$. Let us observe that for this sequence

$$\omega_n(\zeta) = \prod_{\nu=1}^{n} (\zeta - \zeta_\nu^{(n)}) = U_n(\zeta),$$

where $U_n(\zeta)$ are the Tchebysheff-polynomials of second kind; $U_n(\cos \vartheta)$ differs only in a factor from $\sin (n+1)\theta / \sin \theta$, which is independent of ϑ; these polynomials are well known by their many important extremal properties. This

[3] Clearly $0 \leqq \varphi_1^{(n)} < \varphi_2^{(n)} < \cdots < \varphi_n^{(n)} \leqq \pi$.

holds also in the important case $\varphi_\nu^{(n)} = \dfrac{2\nu - 1}{2n}\,\pi$. In both cases exactly $\left[\dfrac{\beta - \alpha}{\pi}\,n\right]\varphi_\nu^{(n)}$ (n fixed), fall in any subinterval $[\alpha, \beta]$ of $[0, \pi]$; the bracket in the last expression means the largest integer contained in $\dfrac{\beta - \alpha}{\pi}\,n$.

According to Weyl's criterion the uniformly dense distribution of a sequence of points is assured by the asymptotic behaviour of certain sequences associated with it. Fekete[4] gives another criterion of the uniform dense distribution; he forms with the sequence of points $\mathfrak{M}_3$ the sequence of polynomials

$$\omega_n(z) = \prod_{\nu=1}^{n} (z - z_\nu^{(n)}) \qquad (n = 1, 2, \cdots)$$

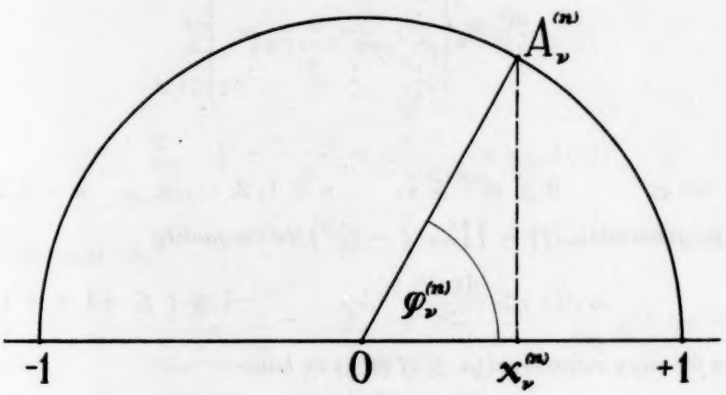

FIG. 1

and shows that $\mathfrak{M}_3$ is uniformly distributed upon l when and only when at every fixed point z_0 of l

$$\overline{\lim_{n\to\infty}}\ (|\,\omega_n(z_0)\,|)^{1/n} \leqq M,$$

where M is the so-called transfinite diameter[5] of l. If l is the interval $[-1\ +1]$ then $M = \frac{1}{2}$; if l is a circle, the radius of which is 1, then $M = 1$. Thus Fekete's criterion requires instead of asymptotic equalities only inequalities.

Let us now consider the special case when l is the interval $[-1, +1]$. In this case we stated that the ideal case is obtained when $\left[\dfrac{\beta - \alpha}{\pi}\,n\right]$ numbers $\varphi_\nu^{(n)}$, (n fixed), fall in any subinterval $[\alpha, \beta]$ of $[0, \pi]$. The above mentioned theorems give no account of the measure of the deviation from this ideal case. In this direction we prove the following

[4] We know this theorem only from an oral communication.

[5] This notion was introduced by M. Fekete: *Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen, etc.*, Math. Zeit., 1923, pp. 228–249.

THEOREM: *Let the sequence of points*

$$\mathfrak{M} \equiv \left\{ \begin{matrix} \zeta_1^{(1)} \\ \vdots \quad \ddots \\ \zeta_1^{(n)}, \zeta_2^{(n)}, \cdots \zeta_n^{(n)} \\ \vdots \quad \vdots \quad \ddots \end{matrix} \right\}$$

with

$$1 \geqq \zeta_1^{(n)} > \zeta_2^{(n)} > \cdots > \zeta_n^{(n)} \geqq -1$$

be given. Let us construct the matrix

$$\mathfrak{M}' \equiv \left\{ \begin{matrix} \varphi_1^{(1)} \\ \vdots \quad \ddots \\ \varphi_1^{(n)}, \varphi_2^{(n)}, \cdots \varphi_n^{(n)} \\ \vdots \quad \vdots \quad \ddots \end{matrix} \right\}$$

with

$$\zeta_\nu^{(n)} = \cos \varphi_\nu^{(n)}, \qquad 0 \leqq \varphi_\nu^{(n)} \leqq \pi, \qquad \nu = 1, 2, \cdots n, \qquad n = 1, 2, \cdots$$

If for the polynomials $\omega_n(\zeta) = \prod_{\nu=1}^{n} (\zeta - \zeta_\nu^{(n)})$ *the inequality*

$$|\omega_n(\zeta)| \leqq \frac{A(n)}{2^n}, \qquad -1 \leqq \zeta \leqq +1, \; n = 1, 2, \cdots, \tag{7}$$

holds, then for every subinterval $[\alpha, \beta]$ *of* $[0, \pi]$ *we have*

$$\left| \sum_{\substack{\nu \\ \alpha \leqq \varphi_\nu^{(n)} \leqq \beta}} 1 - \frac{\beta - \alpha}{\pi} n \right| < \frac{8}{\log 3} (n \log A(n))^{\frac{1}{2}}. \tag{8}$$

Our proof differs thoroughly from that of Fekete and besides it is elementary. $A(n)$ in (7) denotes any function of n tending monotonically to infinity and for which, following Tchebysheff, $A(n) \geqq 2$.

The proof requires a theorem of M. Riesz,[6] the proof of which is so short that for sake of completeness we may reproduce it as a

LEMMA. *Let the trigonometric polynomial* $f(\varphi)$ *of order* n *take its absolute maximum in* $[0, 2\pi]$ *at* $\varphi = \varphi_0$; *then the distance of the next root from this* φ_0 *is at least* $\pi/2n$ *to the right or to the left. Thus a fortiori: if* $f(\varphi)$ *takes its absolute maximum between two real roots, then the distance between these roots is* $\geqq \pi/n$.

PROOF. Suppose that the theorem is false. Without any loss of generality we may assume $\varphi_0 = 0$, there is a maximum at $\varphi = \varphi_0$ and the value of this maximum is 1. Suppose, that the next root lies to the right. Then for $\varphi = 0$

---

[6] M. Riesz: *Eine trigonometrische Interpolations-formel etc.*, Jahresbericht der deutschen Mathematikerver, 1915, Bd. 23, pp. 354–368.

the curves $y = f(\varphi)$ and $y = \cos n\varphi$ would have at least one point of intersection and according to the supposition at least three in $[0, \pi/n]$. There is evidently at least one point of intersection at each of the intervals $[\pi/n, 2\pi/n]$, $[2\pi/n, 3\pi/n]$, $\cdots$ $[(2n-2)\pi/n, (2n-1)\pi/n]$ too; thus the trigonometric polynomial $f(\varphi) - \cos n\varphi$, of order n, would have in $[0, 2\pi]$ $(2n+1)$ roots, an obvious impossibility.

Now we proceed to prove our theorem. It will be sufficient to prove the upper estimate, for if we have proved

$$\sum_{\substack{\nu \\ \alpha \leqq \varphi_\nu^{(n)} \leqq \beta}} 1 - \frac{\beta - \alpha}{\pi} n < \frac{4}{\log 3} (n \log A(n))^{\frac{1}{2}}$$

for every subinterval $[\alpha, \beta]$ between 0 and π, then

$$\text{(9a)} \qquad \sum_{\substack{\nu \\ 0 \leqq \varphi_\nu^{(n)} \leqq \alpha}} 1 - \frac{\alpha}{\pi} n < \frac{4}{\log 3} (n \log A(n))^{\frac{1}{2}}$$

$$\text{(9b)} \qquad \sum_{\substack{\nu \\ \beta \leqq \varphi_\nu^{(n)} \leqq \pi}} 1 - \frac{\pi - \beta}{\pi} n < \frac{4}{\log 3} (n \log A(n))^{\frac{1}{2}}$$

i.e. from (9a) and (9b)

$$\sum_{\substack{\nu \\ \alpha \leqq \varphi_\nu^{(n)} \leqq \beta}} 1 \doteq n - \sum_{\substack{\nu \\ 0 \leqq \varphi_\nu^{(n)} \leqq \alpha}} 1 - \sum_{\substack{\nu \\ \beta \leqq \varphi_\nu^{(n)} \leqq \pi}} 1 > \frac{\beta - \alpha}{\pi} - \frac{8}{\log 3} (n \log A(n))^{\frac{1}{2}},$$

which implies the lower estimate.

Let us now consider the upper estimate. Let $[\alpha, \beta]$ be any subinterval of $[0, \pi]$, but now we regard it as fixed. Let $k = \left[\frac{\beta - \alpha}{\pi} n\right]$, let l be any positive integer, and consider the following extremum problem of the Tchebysheff type: determine the minimum of the absolute maxima of the polynomials $f(\zeta) = \zeta^n + a_1\zeta^{n-1} + \cdots + a_n$ taken in $[-1, +1]$ with the restriction, that $f(\zeta)$ has in the interval $[\cos\beta, \cos\alpha] \equiv [a, b]$, $k + 2l$ roots, (counted by their multiplicity), where as a matter of fact $k + 2l \leqq n$. By a well known argument the existence of this polynomial is assured. We shall prove that the extremum polynomial $f_1(\zeta)$ takes its absolute maximum value with respect to $[-1, +1]$ in each of its root-intervals $[\zeta_{\nu+1}^{(n)}, \zeta_\nu^{(n)}]$ which is in the interior of $[a, b]$. For let the absolute value of the maximum of $f_1(\zeta)$ in $[-1, +1]$ be M and, in $[\zeta_{\nu+2}^{(n)}, \zeta_\nu^{(n)}]$, $|f_1(\zeta)| \leqq M - \eta$, with $\eta > 0$. Then according to a fundamental theorem, for the polynomial

$$f_1^+(\zeta) \equiv \frac{f_1(\zeta)}{(\zeta - \zeta_\nu^{(n)})(\zeta - \zeta_{\nu+1}^{(n)})} (\zeta - \zeta_\nu^{(n)} - \epsilon)(\zeta - \zeta_{\nu+1}^{(n)} + \epsilon)$$

(where $\epsilon > 0$), we have in $[\zeta_{\nu+1}^{(n)} - \epsilon^{\frac{1}{3}}, \zeta_{\nu}^{(n)} + \epsilon^{\frac{1}{3}}]$ $|f_1^+(\zeta)| \leqq M - \frac{1}{2}\eta$, if ϵ sufficiently small. On the other hand in the exterior of $[\zeta_{\nu+1}^{(n)} - \epsilon^{\frac{1}{3}}, \zeta_{\nu}^{(n)} + \epsilon^{\frac{1}{3}}]$, since

$$f_1^+(\zeta) = f_1(\zeta)\left(1 - \frac{\epsilon}{\zeta - \zeta_{\nu}^{(n)}}\right)\left(1 + \frac{\epsilon}{\zeta - \zeta_{\nu+1}^{(n)}}\right)$$

$$= f_1(\zeta)\left\{1 - \frac{\epsilon(\zeta_{\nu}^{(n)} - \zeta_{\nu+1}^{(n)})}{(\zeta - \zeta_{\nu}^{(n)})(\zeta - \zeta_{\nu+1}^{(n)})} - \frac{\epsilon^2}{(\zeta - \zeta_{\nu}^{(n)})(\zeta - \zeta_{\nu+1}^{(n)})}\right\},$$

we have, for $\epsilon < \dfrac{\zeta_{\nu}^{(n)} - \zeta_{\nu+1}^{(n)}}{2}$,

$$|f_1^+(\zeta)| < |f_1(\zeta)|\left|1 - \frac{\epsilon}{2}\frac{\zeta_{\nu}^{(n)} - \zeta_{\nu+1}^{(n)}}{|\zeta - \zeta_{\nu}^{(n)}||\zeta - \zeta_{\nu+1}^{(n)}|}\right| < M\left|1 - \frac{\epsilon^2}{2}(\zeta_{\nu}^{(n)} - \zeta_{\nu+1}^{(n)})\right|.$$

This is less than M for sufficiently small ϵ, which contradicts the fact that $f_1(\zeta)$ is an extremum polynomial. Thus $f_1(\cos\vartheta)$ is a polynomial of order n the $k + 2l$ roots of which in $[\alpha, \beta]$ determine intervals such that in each of these intervals $f(\cos\vartheta)$ takes its absolute maximum.

Now we apply the theorem of M. Riesz formulated in the Lemma. According to this $f(\cos\vartheta)$ cannot have in the interior of $[\alpha, \beta]$ more than $\left[\dfrac{\beta - \alpha}{\pi}n\right] = k$ roots i.e. $2l$ roots must be located at the borders and consequently their multiplicity must be l at least at one of the borders.

By the premise, by the definition of $f_1(\zeta)$, and by the above, we obtain that if $\omega_n(\zeta)$ has $k + 2l$ roots in $[a, b]$, then

$$\frac{A(n)}{2^n} \geqq \max_{-1\leqq\zeta\leqq+1} |\omega_n(\zeta)| \geqq \max_{-1\leqq\zeta\leqq+1} |f_1(\zeta)| \geqq \max_{-1\leqq\zeta\leqq+1} |\psi_1(\zeta)|, \tag{10}$$

where $\psi_1(\zeta) = \zeta^n + b_1\zeta^{n+1} + \cdots + b_n$ denotes the polynomial of degree n the absolute maximum of which is a minimum for polynomials of degree n having *somewhere* in $[-1, +1]$ one root with multiplicity l. As

$$\max_{-1\leqq\zeta\leqq+1} |\psi_1(\zeta)| \geqq \left(\frac{1}{\pi}\int_{-1}^{1} \frac{|\psi_1(\zeta)|^2}{(1 - \zeta^2)^{\frac{1}{2}}}\,d\zeta\right)^{\frac{1}{2}}$$

we have, evidently, (by (10)),

$$\frac{A(n)}{2^n} \geqq \min_{\psi_2}\left(\frac{1}{\pi}\int_{-1}^{1} \frac{|\psi_2(\zeta)|^2}{(1 - \zeta^2)^{\frac{1}{2}}}\,d\zeta\right)^{\frac{1}{2}},$$

where $\psi_2(\zeta) = \zeta^n + \cdots$ runs over the polynomials of order n having somewhere in $[-1, +1]$ a root of the multiplicity l. Let $I_n(\zeta_0)$ denote the minimum value of $\displaystyle\int_{-1}^{1} \frac{|\psi_3(\zeta)|^2}{(1 - \zeta^2)^{\frac{1}{2}}}\,d\zeta$, if $\psi_3(\zeta) = \zeta^n + \cdots$ runs over the polynomials of degree n having *at a fixed* ζ_0 in $[-1, +1]$ a root of multiplicity l. Then we have

$$\frac{A(n)}{2^n} \geqq \min_{\substack{\psi_3 \\ -1\leqq\zeta\leqq+1}}\left(\frac{1}{\pi} I_n(\zeta_0)\right)^{\frac{1}{2}}. \tag{11}$$

Let us now consider $I_n(\zeta_0)$. Every $\psi_3(\zeta)$ can be written in the form $(\zeta - \zeta_0)^l \psi_4(\zeta)$, where $\psi_4(\zeta) = \zeta^{n-l} + \cdots$. Thus

$$I_n(\zeta_0) = \min_{\psi_4(\zeta)=\zeta^{n-l}+\cdots} \int_{-1}^{1} \frac{|\psi_4(\zeta)|^2 |\zeta - \zeta_0|^{2l}}{(1-\zeta^2)^{\frac{1}{2}}} d\zeta.$$

Let $\zeta = \frac{1}{2}\left(z + \frac{1}{z}\right)$, which transforms $[-1, +1]$ into the upper part of the unit circle of the z-plane. Then—for $z = e^{i\varphi}$—

$$I_n(\zeta_0) = \min_{\psi_4(\zeta)=\zeta^{n-l}+\cdots} \frac{1}{2^{2l}} \int_0^{\pi} \left|\psi_4\left(\frac{z+\frac{1}{z}}{2}\right)\right|^2 \left|z + \frac{1}{z} - 2\zeta_0\right|^{2l} d\varphi$$

$$= \min_{\psi_4(\zeta)=\zeta^{n-l}+\cdots} \frac{1}{2^{2l+1}} \int_0^{2\pi} \left|\psi_4\left(\frac{z+\frac{1}{z}}{2}\right)\right|^2 \left|z + \frac{1}{z} - 2\zeta_0\right|^{2l} d\varphi$$

as $|z| = 1$, we have evidently for $\zeta_0 = \cos\alpha_0$

$$I_n(\zeta_0) = \frac{1}{2^{2n+1}} \min_{\psi_5(z)=z^{2n-2l}+\cdots} \int_0^{2\pi} |\psi_5(z)|^2 |z - e^{i\alpha_0}|^{2l} |z - e^{-i\alpha_0}|^{2l} d\varphi$$

$$= \frac{1}{2^{2n+1}} \min \int_0^{2\pi} |\psi_6(z)|^2 d\varphi,$$

where the last minimum is to be taken amongst the polynomials $\psi_6(z) = z^{2n} + \cdots$ of degree $2n$ having at $z = e^{i\alpha_0}$ and $z = e^{-i\alpha_0}$ roots of multiplicity l. Thus

$$I_n(\zeta_0) > \frac{1}{2^{2n+1}} \min_{\psi_7} \int_{|z|=1} |\psi_7(z)|^2 d\varphi,$$

where the minimum relates to the polynomials of degree $2n$ $\psi_7(z) = z^{2n} + \cdots$ having an l-fold root only at $z = e^{i\alpha_0}$. But in this case, since $\psi_7(z) = (z - e^{i\alpha_0})^l \psi_8(z)$, we have

$$I_n(\zeta_0) > \frac{1}{2^{2n+1}} \min_{\psi_8} \int_{|z|=1} |\psi_8(z)|^2 |z - e^{i\alpha_0}|^{2l} d\varphi,$$

where $\psi_8(z)$ runs over the polynomials of degree $(2n - l)$ beginning with z^{2n-l}. Finally by replacing φ by $\alpha_0 + \varphi + \pi$ we obtain

$$I_n(\zeta_0) > \frac{1}{2^{2n+1}} \min \int_{|z|=1} |\psi_9(z)|^2 |1 + z|^{2l} d\varphi, \tag{12}$$

the minimum being taken for all polynomials $\psi_9(z) = z^{2n-l} + \cdots$ of degree $(2n - l)$.

If $p(\varphi)$ defines in $[0, 2\pi]$ a non negative and L-integrable function then, after Szegö we may define a sequence of polynomials $\phi_0(z)$, $\phi_1(z)$, $\cdots$ such that

$$\phi_m(z) = z^m + \cdots \qquad m = 1, 2, \cdots \tag{13a}$$

$$\int_{|z|=1} \phi_m(z)(\bar{z})^\nu p(\varphi)\, d\varphi = 0 \qquad \begin{matrix} \nu = 0, 1, \cdots, (m-1); \\ m = 1, 2, \cdots. \end{matrix} \tag{13b}$$

In this case $\phi_m(z)$ minimizes, for polynomials $U(z)$ of degree m of the form $U(z) = z^m + \cdots$, the integral $\int_{|z|=1} |U(z)|^2 p(\varphi)\, d\varphi$. For any other such polynomial may be reduced the form $\phi_m(z) + \pi_{m-1}(z)$, where $\pi_{m-1}(z)$ is a polynomial of degree $(m-1)$. Then by (13b)

$$\int_{|z|=1} |\phi_m(z) + \pi_{m-1}(z)|^2 p(\varphi)\, d\varphi$$

$$= \int_{|z|=1} |\phi_m(z)|^2 p(\varphi)\, d\varphi + 2\Re \int_{|z|=1} \phi_m(z)\overline{\pi_{m-1}(z)} p(\varphi)\, d\varphi + \int_{|z|=1} |\pi_{m-1}(z)|^2 p(\varphi)\, d\varphi$$

$$= \int_{|z|=1} |\phi_m(z)|^2 p(\varphi)\, d\varphi + \int_{|z|=1} |\pi_{m-1}(z)|^2 p(\varphi)\, d\varphi \geqq \int_{|z|=1} |\phi_m(z)|^2 p(\varphi)\, d\varphi,$$

and equality holds only for $\pi_{m-1}(z) \equiv 0$. The expression on the right side of (12) takes its minimum value for the $(2n-l)^{\text{th}}$ polynomial orthogonal to the weight-function $p(\varphi) = |1+z|^{2l}$. According to a theorem of Szegö these polynomials may be expressed in terms of Jacobi polynomials but we prefer to present them in the form of an explicit integral interesting in itself. For $m = 2n - l$ we write

$$(14)\quad \phi_{2n-l}(z) = \frac{l\binom{2n+l}{l}}{(1+z)^{2l}} \int_{-1}^{z} (z-t)^{l-1}(1+t)^l t^{2n-l}\, dt \equiv \frac{l\binom{2n+l}{l}}{(1+z)^{2l}} F_{2n+l}(z).$$

This expression is a polynomial; we prove it by showing that $F_{2n+l}(-1) = F'_{2n+l}(-1) = \cdots = F^{(2l-1)}_{2n+l}(-1) = 0$. The first of these equations is an immediate consequence of (14). Since for $1 \leqq \nu \leqq l-1$ we have by (14)

$$F^{(\nu)}_{2n+l}(z) = (l-1)(l-2)\cdots(l-\nu) \int_{-1}^{z} (z-t)^{l-\nu-1}(1+t)^l t^{2n-l}\, dt,$$

it is evident that the assertion holds for $1 \leqq \nu \leqq l-1$. On the other hand

$$F^{(l)}_{2n+l}(z) = (l-1)!\,(1+z)^l z^{2n-l},$$

i.e. evidently $F^{(l)}_{2n+l}(-1) = \cdots = F^{(2l-1)}_{2n+l}(-1) = 0$. Hence the expression in (14) is a polynomial.

We now prove that the coefficient of z^{2n-l} in (14) equals 1. The coefficient in question is

$$l\binom{2n+l}{l} \lim_{z\to\infty} \frac{\int_{-1}^{z} (z-t)^{l-1}(1+t)^l t^{2n-l}\, dt}{z^{2n+l}}$$

$$= l\binom{2n+l}{l} \lim_{z\to\infty} \frac{\int_{-1}^{z} (z-t)^{l-1} t^{2n}\, dt}{z^{2n+l}},$$

which by the substitution $t = zw$ can be transformed into

$$l\binom{2n+l}{l} \lim_{z\to\infty} \int_{-1/z}^{\infty} (1-w)^{l-1} w^{2n}\, dw = l\binom{2n+l}{l} \int_0^{\infty} (1-w)^{l-1} w^{2n}\, dw = 1.$$

And now we have to verify the relation of orthogonality. Let

$$\int_{|z|=1} \phi_{2n-l}(z)(\bar{z})^{\nu} |1+z|^{2l}\, d\varphi \equiv A_{\nu}. \qquad \nu = 0, 1, \cdots (2n-l-1).$$

Since, for $|z| = 1$,

$$|1+z|^{2l} = \frac{(1+z)^{2l}}{z^l}, \qquad \bar{z} = \frac{1}{z},$$

we have by (14)

$$A_{\nu} = l\binom{2n+l}{l} \int_{|z|=1} \frac{F_{2n+l}(z)}{z^{\nu+l}}\, d\varphi = \frac{l\binom{2n+l}{l}}{i} \int_{|z|=1} \frac{F_{2n+l}(z)}{z^{\nu+l+1}}\, dz.$$

Hence if in $F_{2n+l}(z)$ the coefficient of $z^l, z^{l+1}, \cdots, z^{2n-l}$ equals 0, (13b) is verified. But according to the definition of $F_{2n+l}(z)$

$$(15) \quad F_{2n+l}(z) = \int_{-1}^{0} (z-t)^{l-1}(1+t)^l t^{2n-l}\, dt + \int_0^z (z-t)^{l-1}(1+t)^l t^{2n-l}\, dt.$$

Here the first integral is a polynomial of z of degree $(l-1)$; thus it has no influence upon our assertion. The second one we transform by the substitution $t = zw$ into

$$z^{2n} \int_0^1 (1-w)^{l-1}(1+2w)^l w^{2n-l}\, dw.$$

Thus the second term is a polynomial the *lowest* term of which is $2n$, which establishes the orthogonality.

The minimum value is given by

$$\int_{|z|=1} |\phi_{2n-l}(z)|^2 |1+z|^{2l}\, d\varphi = \int_{|z|=1} \phi_{2n-l}(z)\{(\bar{z})^{2n-l} + \cdots\} |1+z|^{2l}\, d\varphi$$

$$= \int_{|z|=1} \phi_{2n-l}(z)(\bar{z})^{2n-l} |1+z|^{2l}\, d\varphi = \frac{l\binom{2n+l}{l}}{i} \int_{|z|=1} \frac{F_{2n+l}(z)}{z^{2n+1}}\, dz$$

$$= 2\pi l\binom{2n+l}{l} \times (\text{the coefficient of } z^{2n} \text{ in } F_{2n+l}(z)),$$

which by the form of $F_{2n+l}(z)$ in (15) equals

$$2\pi\binom{2n+l}{l} \int_0^1 (1-w)^{l-1} w^{2n-l}\, dw = 2\pi \frac{\binom{2n+l}{l}}{\binom{2n}{l}}.$$

By this and (12)

$$I_n(\zeta_0) > \frac{\pi}{2^{2n}} \frac{\binom{2n+l}{l}}{\binom{2n}{l}},$$

i.e., by (11),

$$\frac{A(n)}{2^n} > \frac{1}{2^n}\left(\frac{\binom{2n+l}{l}}{\binom{2n}{l}}\right)^{\frac{1}{2}} = \frac{1}{2^n}\left(\prod_{\nu=0}^{l-1}\left(1+\frac{l}{2n-\nu}\right)\right)^{\frac{1}{2}}. \tag{16}$$

Since $\frac{l}{2n-\nu} \leqq 2$ and $\log(1+x) \geqq \frac{\log 3}{2}x$, (if $0 \leqq x \leqq 2$), we have, (by (16)),

$$A(n) > \operatorname{Exp}\left[\frac{\log 3}{4} l \sum_{\nu=2n-l+1}^{2n} \frac{1}{\nu}\right]$$

$$> \operatorname{Exp}\left[\frac{\log 3}{4} l \log\left(1+\frac{l}{2n-l+1}\right)\right] > \operatorname{Exp}\left[\left(\frac{\log 3}{4}\right)^2 \frac{l^2}{n}\right],$$

$$l < \frac{4}{\log 3}(n \log A(n))^{\frac{1}{2}},$$

which establishes the result.

NOTE I. For the Tchebysheff-polynomial $T_n(x)$, where $T_n(\cos\vartheta) = \frac{\cos n\vartheta}{2^{n-1}} = \cos^n\vartheta + \cdots$, we have in $[-1, +1]$ evidently $|T_n(x)| \leqq \frac{1}{2^{n-1}}$, i.e. $T_n(x)$ approximates the function $y \equiv 0$ in Tchebysheff's sense, the error being less then $\frac{2}{2^n}$. By the above argument it can be seen that the function $y \equiv 0$ is to be approximated not essentially worse, in Bessel's sense, by a polynomial of the form $x^n + \cdots$, even when the polynomial has somewhere in $[-1, +1]$ a root the multiplicity of which is less than $[\sqrt{n}]$. We are of the opinion that this very probably holds also for the Tchebysheff approximation; i.e. there exists a polynomial of degree n $f(x) = x^n + \cdots$, which has somewhere in $[-1, +1]$ a root of the multiplicity $[\sqrt{n}]$, and yet in $[-1, +1]$

$$|2^n f(x)| < B,$$

where B is independent of n. By this it is clear that in general, the above theorem is not to be improved.

NOTE II. Let $\omega_n(x) = x^n + \cdots$ be the polynomial of degree n minimizing for polynomials of degree n of the form $f(x) = x^n + \cdots$ the integral

$$I_k(f) \equiv \int_{-1}^{1} |f(x)|^k p(x)\,dx, \tag{17}$$

where k is a fixed positive number, $p(x)$ is L-integrable, and in $[-1, +1] p(x) \geqq m(> 0)$. According to a theorem of Fejér all roots of $\omega_n(x)$ are in $[-1, +1]$. Denote the absolute maximum of $|\omega_n(x)|$ by M; if this maximum is taken at $x = x_0$ and one of the intervals $\left[x_0 - \frac{1}{2n^2}, x_0\right], \left[x_0, x_0 + \frac{1}{2n^2}\right]$, suppose the latter, lies in $[-1, +1]$, then

$$\int_{-1}^{1} |\omega_n(x)|^k p(x)\, dx > m \int_{x_0}^{x_0 + \frac{1}{2}n^{-2}} |\omega_n(x)|^k\, dx.$$

But by Markoff's theorem $|\omega_n(x)| > \frac{M}{2}$ in $\left[x_0, x_0 + \frac{1}{2n^2}\right]$, i.e.

$$\int_{-1}^{1} |\omega_n(x)|^k p(x)\, dx > m \left(\frac{M}{2}\right)^k \frac{1}{2n^2}. \tag{18}$$

On the other hand, if $f(x) = T_n(\cos \vartheta) = \frac{\cos n\vartheta}{2^{n-1}}$, we obtain by the minimum-property of $\omega_n(x)$

$$\int_{-1}^{1} |\omega_n(x)|^k p(x)\, dx < \frac{1}{2^{k(n-1)}} \int_{-1}^{1} p(x)\, dx, \tag{19}$$

i.e. by (18) and (19), for $[-1, +1]$ we have

$$|\omega_n(x)| \leqq M < \frac{2^{2+k-1}}{m^{k-1}} \left(\int_{-1}^{1} p(t)\, dt\right)^{k-1} \frac{n^{2k-1}}{2^n}.$$

Thus by the theorem mentioned above we obtain for the roots of polynomials minimizing the expressions in (17) that if the roots on the n^{th} polynomials are $x_1^{(n)}, x_2^{(n)} \cdots, x_n^{(n)}$, and $x_\nu^{(n)} = \cos \vartheta_\nu^{(n)}$, $0 \leqq \vartheta_\nu^{(n)} \leqq \pi$, then for any fixed subinterval $[\alpha, \beta]$

$$\left| \sum_{\alpha \leqq \vartheta_\nu^{(n)} \leqq \beta} 1 - \frac{\beta - \alpha}{\pi} n \right| < c(p, k)(n \log n)^{\frac{1}{2}}$$

i.e., roughly speaking, the distribution of the roots of the minimizing polynomials is uniformly dense. Analogous theorems are to be deduced for the polynomials solving extremum problems of the Tchebysheff-type.

INSTITUTE FOR ADVANCED STUDY, PRINCETON, N. J.
AND BUDAPEST VI., HUNGARY.

ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

SUR QUELQUES NOTIONS FONDAMENTALES DANS LA THÉORIE GÉNÉRALE DES OPÉRATIONS LINÉAIRES

PAR FRÉDÉRIC RIESZ

(Received April 8, 1939)

Le présent Mémoire a pour but de développer et de poursuivre quelques idées concernant les opérations linéaires, idées que j'ai soulevées dans une conférence faite en 1928, au congrès international des mathématiciens à Bologne et dont je me suis servi pour décomposer ces opérations en des parties appartenant à des types particuliers.[1] Là je me suis borné, à titre d'exemple, à des opérations linéaires portant sur des fonctions continues et j'ai montré, entre autres, que les parties en question de l'opération linéaire peuvent être caractérisées et obtenues sans avoir recours à l'expression analytique de l'opération sous forme d'intégrale de Stieltjes. Comme je l'ai déjà dit dans ma conférence, une telle méthode a l'avantage d'être entièrement générale de sorte que l'on peut s'en servir pour les opérations linéaires des fonctions définies dans des ensembles abstraits, comme les a étudiées M. Daniell en généralisant la notion d'intégrale.[2] Les hypothèses dont nous partirons dans le présent Mémoire, sont encore plus générales que celles auxquelles j'ai pensé au temps de ma conférence; ce n'est pas à l'intervalle ou aux ensembles de points que nous substituerons des ensembles abstraits, ni même les fonctions continues remplacées par une classe plus générale de fonctions, mais ce sont les fonctions elles-mêmes dont le rôle sera joué par des éléments abstraits, d'ailleurs soumis à quelques très peu d'hypothèses concernant leur addition.

Notre théorie est en relation intime avec celle des modules semiordonnés et des lattis linéaires, récemment développée par MM. Kantorovitch, Freudenthal et Garrett Birkhoff.[3] La différence la plus essentielle entre les deux théories concerne leurs points de départ, la théorie des lattis étant fondée sur des hypothèses plus exigeantes que celles posées pour notre domaine fondamental, c'est l'ensemble de nos éléments abstraits. Malgré cela, les opérations linéaires que l'on peut définir dans ce domaine forment ce qu'on appelle un lattis linéaire complet, c'est-à-dire que pour l'ensemble de ces opérations, les hypothèses de l'autre théorie s'y réalisent d'elles-mêmes. Tel est aussi le cas pour le pro-

[1] F. Riesz, "Sur la décomposition des opérations linéaires," *Atti del congresso internazionale dei matematici*, Bologna 1928, **3**, pp. 143–148.

[2] P. J. Daniell, "The Integral and Its Generalizations," *Rice Institute Pamphlet*, **8** (1921), pp. 34–62; "A General Form of Integral," *Annals of Math.*, (2) **29** (1917–18), pp. 279–294; "Further Properties of the General Integral," *ibidem*, (2) **31** (1919–20), pp. 203–220.

[3] Pour le présent travail *cf.* en premier lieu: H. Freudenthal, "Teilweise geordnete Moduln," *Proceedings Acad. Amsterdam*, **39** (1936), pp. 641–651. L'auteur date l'origine de la théorie en question de ma conférence de Bologne.

longement de notre domaine fondamental qui se fait par une sorte de dualité, en échangeant le rôle des éléments et des opérations. En révanche, la théorie des lattis qui est encore plus abstraite, pourra nous aider d'examiner de plus proche les rapports mutuels qui existent, au point de vue formel, entre nos résultats concernant les opérations linéaires.

C'est en 1936 que j'ai achevé et présenté à l'Académie hongroise une première rédaction, en langue hongroise, du présent Mémoire.[4] La rédaction actuelle en diffère, à part quelques modifications légères ici et là, principalement en ce qu'elle comprend un chapitre consacré au calcul fonctionnel des opérations linéaires, calcul permettant de composer des opérations données d'une manière très générale. Quelques compositions de type particulier, comme les produits, les puissances et les moyennes géométriques sont traitées déjà dans la première édition par une méthode entièrement différente qui n'est pas applicable au cas général. Aussi, pour épargner d'interrompre l'ordre d'idées des divers chapitres, nous avons intercalé quelques théorèmes plutôt auxiliaires. Enfin nous avions à refaire le dernier chapitre qui rend compte de quelques idées vagues concernant le prolongement du domaine fondamental et dans lequel se sont glissées quelques indications erronées.

Observons encore, ce qui est d'ailleurs manifeste, que dans les considérations qui suivent, on aurait pu remplacer les opérations numériques par d'autres dont les valeurs appartiennent à n'importe quel lattis linéaire complet.

I. Le domaine fondamental; exemples

Nous allons envisager des opérations définies dans un ensemble abstrait que nous appelons le domaine fondamental et que nous désignons par Ω. Les applications les plus importantes de notre théorie étant celles relatives à des opérations fonctionnelles, nous désignerons les éléments de Ω en général par les lettres f et g. Nous supposons définie pour le domaine Ω une addition faisant correspondre aux éléments f_1 et f_2 un élément $f_1 + f_2$ bien déterminé, correspondance qui aura à satisfaire aux hypothèses suivantes.

1. *Elle est commutative et associative*: $f_1 + f_2 = f_2 + f_1$; $(f_1 + f_2) + f_3 = f_1 + (f_2 + f_3)$.

2. *La relation* $f_1 + f_2 = f_1 + f_3$ *implique que* $f_2 = f_3$.

3. *Il y a un élément* 0, *tel que* $f + 0 = f$ *pour tous les* f *et la relation* $f_1 + f_2 = 0$ *implique que* $f_1 = f_2 = 0$.

4. *Lorsque* $f_1 + f_2 = g_1 + g_2$, *il existe quatre éléments* f_{11} , f_{12} , f_{21} , f_{22} *de sorte que* $f_{11} + f_{12} = f_1$, $f_{21} + f_{22} = f_2$, $f_{11} + f_{21} = g_1$, $f_{12} + f_{22} = g_2$.

Voici une forme plus générale de cette dernière hypothèse qui s'en déduit par récurrence, ayant égard encore à l'hypothèse 1. *Étant donnés des éléments* f_i , g_k $(i = 1, \cdots, m;\ k = 1, \cdots, n)$ *tels que*

$$\sum_1^m f_i = \sum_1^n g_k,$$

[4] Imprimée dans les comptes rendus de la Classe des Sciences *Matematikai és Természettudományi Értesitő*, **56** (1937), pp. 1–46.

il existe des éléments f_{ik} de sorte que

$$\sum_{k=1}^{n} f_{ik} = f_i; \qquad \sum_{i=1}^{m} f_{ik} = g_k.$$

Considérons quelques exemples. Les plus simples sont l'ensemble des nombres réels non-négatifs, celui des entiers non-négatifs et celui des logarithmes des entiers positifs ou ce qui revient au même, l'ensemble des entiers positifs, mais en remplaçant l'addition par la multiplication et le zéro par l'unité. Outre ces exemples évidents, en voici quelques autres, plus intéressants au point de vue de la théorie que nous allons développer. Dans tous ces exemples, l'addition sera définie par l'addition numérique. Les trois premières hypothèses étant évidemment remplies, il ne s'agira que de vérifier l'hypothèse 4.

EXEMPLE 1. Les fonctions non-négatives et continues définies dans l'intervalle $a \leqq x \leqq b$.

Lorsque $f_1 + f_2 = g_1 + g_2 = f$, la décomposition exigée par l'hypothèse 4 est fournie par

$$f_{ik}(x) = \begin{cases} \dfrac{f_i(x)g_k(x)}{f(x)} & \text{quand } f(x) \neq 0, \\ 0 & \text{quand } f(x) = 0. \end{cases}$$

Les relations exigées étant évidentes, de même que la continuité des fonctions f_{ik} aux points x où $f(x)$ ne s'annule pas, nous n'avons qu'à examiner l'allure de ces fonctions aux zéros de la fonction $f(x)$. Or soit x_0 un tel point; alors $f_i(x_0) = 0$ et comme d'autre part $g_k(x) \leqq f(x)$, il vient que $f_{ik}(x) \leqq f_i(x)$ et par conséquent $f_{ik}(x)$ s'annule au point x_0 d'une façon continue.

Voici encore un second procédé de décomposition. Posons

$$f_{11}(x) = \min\,(f_1(x), g_1(x)),$$

$$f_{12} = f_1 - f_{11}, \qquad f_{21} = g_1 - f_{11}, \qquad f_{22} = f - \max\,(f_1(x), g_1(x)).^5$$

Alors il est manifeste que

$$f_{11} + f_{12} = f_1, \qquad f_{11} + f_{21} = g_1$$

et les deux autres relations s'obtiennent en observant que l'on a toujours

$$\min\,(x, y) + \max\,(x, y) = x + y.$$

EXEMPLE 2. Les fonctions rationnelles $f(x) = p(x)/q(x)$ formées de polynomes $p(x) \geqq 0$, $q(x) > 0$ $(a \leqq x \leqq b)$.

La décomposition répondant à l'hypothèse 4 s'effectue par le premier procédé

---

[5] min (x, y) et max (x, y) désignent respectivement la plus petite et la plus grande des valeurs x, y; nous nous servirons de la même notation pour désigner les bornes inférieures et supérieures des ensembles numériques finis ou infinis et cela en réservant les notations inf et sup pour quelques autres catégories d'ensembles.

appliqué dans l'exemple 1. Le second procédé est en défaut, les fonctions des types min (f, g) et max (f, g) n'appartenant pas en général au domaine envisagé.

EXEMPLE 3. Les fonctions escalier non-négatives définies dans l'intervalle $a \leqq x \leqq b$.

EXEMPLE 4. Les mêmes fonctions, mais en convenant de considérer comme identiques deux fonctions quand elles ne diffèrent que pour un nombre fini de valeurs de la variable.

Dans ces deux exemples, la décomposition exigée par l'hypothèse 4 s'effectue par l'un ou l'autre des procédés dont nous nous sommes servis dans l'exemple 1.

EXEMPLE 5. L'ensemble des fonctions harmoniques non-négatives, définies dans une région ouverte du plan ou de l'espace.

La décomposition s'effectue par le second procédé, mais en substituant aux enveloppes min (f_1, g_1) et max (f_1, g_1) qui en général ne seront pas harmoniques, les deux fonctions harmoniques que nous convenons de désigner par $\varphi = \inf(f_1, g_1)$ et $\psi = \sup(f_1, g_1)$ et dont la première est la plus grande parmi les fonctions harmoniques ne surpassant ni f_1 ni g_1 et la seconde se définit d'une manière analogue. Ces deux fonctions peuvent être construites par exemple en appliquant la méthode de balayage de Poincaré aux fonctions min (f_1, g_1) et max (f_1, g_1) dont la première est surharmonique, la seconde sousharmonique. La relation $\varphi + \psi = f_1 + g_1$ est une conséquence immédiate de la relation $\min(f_1, g_1) + \max(f_1, g_1) = f_1 + g_1$, le balayage étant distributif et ne changeant pas le second membre de la dernière relation. Il convient d'ailleurs d'observer que la relation $\varphi + \psi = f_1 + g_1$ se déduit aussi des propriétés extrémales des fonctions φ et ψ. En effet, comme $\varphi \leqq f_1$ et $\leqq g_1$, on a $f_1 + g_1 - \varphi \geqq f_1$ et $\geqq g_1$, donc $f_1 + g_1 - \varphi \geqq \sup(f_1, g_1) = \psi$, ce qui donne

$$f_1 + g_1 \geqq \varphi + \psi. \tag{1}$$

L'inégalité opposée

$$f_1 + g_1 \leqq \varphi + \psi \tag{2}$$

s'obtient en partant de $\psi \geqq f_1$ et $\geqq g_1$, ce qui s'écrit aussi $f_1 + g_1 - \psi \leqq g_1$ et $\leqq f_1$; il s'ensuit que $f_1 + g_1 - \psi \leqq \inf(f_1, g_1) = \varphi$, c.-à-d. l'inégalité (2). La comparaison des inégalités (1) et (2) donne $f_1 + g_1 = \varphi + \psi$, ce qu'il fallait prouver.

Ajoutons que ce dernier raisonnement est valable aussi pour un domaine Ω abstrait. En effet, convenons d'écrire $f_1 \geqq f_2$, $f_2 \leqq f_1$ quand $f_2 = f_1 - f_3$ c'est-à-dire que $f_1 = f_2 + f_3$ et de dire brièvement que f_1 est supérieur à f_2 et f_2 est inférieur à f_1 et de parler conformément à ces conventions des éléments le plus grand et le plus petit d'un ensemble; alors notre raisonnement s'applique et fournit l'énoncé suivant.

Étant donnés deux éléments f et g du domaine Ω, supposons que parmi les éléments qui sont inférieurs à tous les deux il y en a un élément φ qui est le plus grand et que parmi les éléments supérieurs à f et g il y en a un, soit ψ, qui est le plus petit, alors

$$\varphi + \psi = f + g.$$

Si de plus nous désignons par 4′ l'hypothèse que nous venons de formuler, supposée remplie pour tous les f et g, alors l'hypothèse 4 est une conséquence des hypothèses 1–3 et 4′. Au contraire l'hypothèse 4′ n'est pas une conséquence des hypothèses 1–4, comme le montre l'exemple 2.

Exemple 6. L'ensemble des distributions de masses positives dans l'espace, données par exemple sous forme de fonctions additives d'ensemble $\mu(e) \geqq 0$ pour tous les ensembles e ouverts et bornés; on suppose que $\mu(e)$ satisfait aux hypothèses suivantes: 1) $\mu(e') \leqq \mu(e)$ lorsque e' est compris dans e; 2) $\mu(e_1 e_2) + \mu(e_1 + e_2) = \mu(e_1) + \mu(e_2)$; 3) $\mu(e)$ peut être approché indéfiniment par $\mu(e')$, e' parcourant les ensembles ouverts compris dans e de même que leurs frontières.

L'hypothèse 4 est satisfaite aussi pour ce dernier exemple, les distributions inf (μ, ν) et sup (μ, ν) correspondant aux distributions $\mu(e)$ et $\nu(e)$ étant définies, comme on sait, la première par la borne inférieure de $\mu(e_1) + \nu(e_2)$, les ensembles ouverts e_1, e_2 variant sous la condition $e_1 + e_2 = e$, et la seconde par la borne supérieure de $\mu(e_3) + \nu(e_4)$ où e_3 et e_4, sousensembles ouverts de e, varient sous la condition de ne pas s'empiéter.

Exemple 7. Tout système de fonctions $f(x)$ non-négatives définies dans un ensemble abstrait, comprenant la fonction zéro et avec les fonctions $f(x)$, $g(x)$, leur somme $f(x) + g(x)$ et leurs enveloppes min $(f(x), g(x))$ et max $(f(x), g(x))$. C'est à de tels systèmes que s'applique la théorie générale de l'intégration due à M. Daniell.

On arrive à d'autres exemples par des modifications évidentes des exemples que nous venons d'énumérer. Ainsi dans l'exemple 6, on pourrait partir, au lieu de l'espace entier, d'un domaine borné de cet espace ou d'une surface fermée ou d'un intervalle de droite et pour parler aussi de l'autre extrême, d'un espace à une infinité de dimensions ou encore d'un espace abstrait. Dans l'exemple 7 on pourrait convenir, comme nous l'avons déjà fait dans l'exemple 4, de considérer comme identiques deux fonctions qui ne diffèrent que pour un nombre fini de points x. Observons qu'un cas particulier important de l'exemple 7 est fourni par l'ensemble des fonctions non-négatives presque périodiques. Nous pourrions encore ranger dans cet ordre d'idées certaines variétés de transformations linéaires de l'espace de Hilbert, variétés qui ont fait l'objet de travaux importants il y a peu d'années.

II. Opérations linéaires positives; la plus grande minorante et la plus petit majorante

Dans ce qui suit, nous allons considérer des opérations linéaires définies dans le domaine Ω; cependant, jusqu'à la fin du chapitre V, il ne sera question que des opérations linéaires positives. Par là nous entendons les opérations $A = A(f) = Af$ faisant correspondre à chaque élément f une valeur non-négative Af de sorte que $A(f_1 + f_2) = Af_1 + Af_2$.

Lorsque, pour deux opérations, A_1 et A_2, on a $A_1 f \leqq A_2 f$ pour tous les f, nous convenons d'écrire aussi $A_1 \leqq A_2$ et $A_2 \geqq A_1$ et de dire que A_1 est une

minorante de A_2 et que A_2 est une *majorante* de A_1. Un ensemble $\{A\}$ d'opérations A, en nombre fini ou infini, dénombrable ou non, sera dit majorable lorsqu'il existe une opération B telle que $A \leqq B$ pour tous les A. Les opérations B du type envisagé seront dites majorantes de l'ensemble $\{A\}$. D'une manière analogue, lorsque $B \leqq A$ pour tous les éléments de l'ensemble $\{A\}$, on dira que B est une des minorantes de l'ensemble. Tout ensemble d'opérations linéaires positives est minorable, dans un sens évident, une minorante étant toujours fournie par l'opération 0 qui s'annule pour tous les f.

Convenons encore de dire que la majorante B est la plus petite majorante de l'ensemble $\{A\}$ et de la désigner par sup $\{A\}$ lorsque'elle est une minorante de toutes les autres majorantes, ainsi que de définir d'une façon analogue la plus grande minorante et de la désigner par inf $\{A\}$.

THÉORÈME 1. *Chaque ensemble $\{A\}$ d'opérations linéaires positives admet une plus grande minorante, et chaque ensemble $\{A\}$ majorable admet aussi une plus petite majorante.*

Démontrons d'abord la première assertion. A cet effet, nous envisageons toutes les sommes de la forme

$$\sum_1^m A_i f_i \tag{1}$$

où m est arbitraire, où les A_i sont tirés arbitrairement de l'ensemble $\{A\}$ et où les

$$f = \sum_1^m f_i$$

sont des décompositions de l'élément donné f. Variée sous ces conditions, l'expression (1) admet une borne inférieure Bf et cela pour tous les f. Je dis que cette borne Bf, envisagée comme fonction de f, définit l'opération inf $\{A\}$ exigée par notre théorème. Tout d'abord, il est manifeste que $Bf \geqq 0$ et plus généralement, comme

$$\sum_1^m A_i f_i \geqq \sum_1^m C f_i = Cf$$

pour toute minorante C de l'ensemble $\{A\}$, on a aussi $Bf \geqq Cf$ pour toutes ces minorantes. D'autre part on a $Bf \leqq Af$ pour toute opération A; il suffit de poser $m = 1, A_1 = A, f_1 = f$. Donc il nous reste seulement à vérifier la relation

$$B(f + g) = Bf + Bg. \tag{2}$$

A cet effet, observons d'une part que, toute décomposition $f = \sum_1^m f_i$, $g = \sum_1^n g_i$ donnant lieu à une décomposition de $f + g$ en $\sum_1^m f_i + \sum_1^n g_i$, on a immédiatement que

$$B(f + g) \leqq Bf + Bg. \tag{3}$$

D'autre part, soit $f + g = \sum_1^p f_i'$ une décomposition de $f + g$; alors, par l'hypothèse 4 faite concernant le domaine Ω, il existe, comme nous l'avons observé, des décompositions

$$f = \sum_1^p f_i, \qquad g = \sum_1^p g_i$$

telles que $f_i + g_i = f_i'$ $(i = 1, \cdots, p)$. Or, $B(f + g)$ étant la borne inférieure de

$$\sum_1^p A_i f_i' = \sum_1^p A_i f_i + \sum_1^p A_i g_i \geqq Bf + Bg$$

il faut que

$$B(f + g) \geqq Bf + Bg$$

ce qui, comparé avec (3), donne la relation (2) qu'il nous fallait prouver.

Quant à l'existence de la plus grande majorante, ce problème se ramène au précédant et cela en considérant l'ensemble $\{D_0 - A\}$ où A parcourt l'ensemble $\{A\}$ et D_0 désigne n'importe quelle majorante de ce dernier. Je dis que l'opération

$$E = D_0 - \inf \{D_0 - A\} \tag{4}$$

fournit l'opération $\sup \{A\}$. En effet, on a évidemment

$$E \geqq D_0 - (D_0 - A) = A$$

pour tous les A; d'autre part, D_1 désignant l'une quelconque des majorantes de l'ensemble $\{A\}$, l'opération $D_2 = \inf (D_0, D_1)$ sera une majorante du même ensemble et par conséquent,

$$E = D_0 - \inf \{D_0 - A\} \leqq D_0 - (D_0 - D_2) = D_2 \leqq D_1$$

c'est-à-dire que E est bien la plus petite parmi les majorantes de l'ensemble $\{A\}$.

La construction de l'opération $\sup \{A\} = E$ montre aussi immédiatement que Ef est la borne supérieure des valeurs de (1).

Voici encore deux relations, conséquences immédiates des raisonnement qui précèdent; comme nous aurons à nous y reporter à plusieurs reprises, nous les énonçons sous forme de théorèmes.

THÉORÈME 2. *Pour deux opérations linéaires positives A_1, A_2 on a toujours*

$$\inf (A_1, A_2) + \sup (A_1, A_2) = A_1 + A_2. \tag{5}$$

THÉORÈME 3. *Étant donné un ensemble $\{A\}$ d'opérations linéaires positives A et en désignant par $\{A + B_0\}$ l'ensemble obtenu en ajoutant à chaque A l'opération linéaire positive B_0, toujours la même, on a*

$$\inf \{A + B_0\} = \inf \{A\} + B_0$$

et quand l'ensemble $\{A\}$ est majorable, on a aussi

$$\sup \{A + B_0\} = \sup \{A\} + B_0.$$

D'une façon plus générale, lorsque $\{A\}$ *et* $\{B\}$ *sont deux ensembles d'opérations linéaires positives et* $\{A + B\}$ *l'ensemble de toutes les opérations* $A + B$ *que l'on en peut composer, alors on a*

$$\inf \{A + B\} = \inf \{A\} + \inf \{B\} \tag{6}$$

et quand les deux ensembles sont supposés majorables, on a encore

$$\sup \{A + B\} = \sup \{A\} + \sup \{B\}.$$

Le théorème 2 se déduit immédiatement de la formule (4) en y posant $D_0 = A_1 + A_2$, ce qui donne la relation

$$\sup (A_1, A_2) = E = A_1 + A_2 - \inf (A_2, A_1),$$

équivalente à la relation (5).

La première partie du théorème 3 se déduit du fait que les opérations $\inf \{A + B_0\}$ et $\sup \{A + B_0\}$ se calculent comme $\inf \{A\}$ et $\sup \{A\}$, en remplaçant l'expression (1) par

$$\sum_1^m (A_i + B_0) f_i = \sum_1^m A_i f_i + B_0 f.$$

La seconde partie se ramène à la première en partant du fait évident que $\inf \{A + B\} \leqq \inf \{A + B_0\}$ et $\inf \{A\} \leqq A_0$, $\inf \{B\} \leqq B_0$ pour chaque choix particulier des opérations A_0, B_0 appartenant respectivement aux ensembles $\{A\}$, $\{B\}$. En effet, l'inégalité

$$\inf \{A + B\} - \inf \{A\} \leqq \inf \{A + B_0\} - \inf \{A\} = B_0,$$

vraie pour toute opération B_0 tirée de $\{B\}$, nous donne

$$\inf \{A + B\} - \inf \{A\} \leqq \inf \{B\}. \tag{7}$$

D'autre part on a

$$\inf \{A\} + \inf \{B\} \leqq A_0 + B_0$$

pour tout choix particulier des opérations A_0, B_0 et cela donne

$$\inf \{A\} + \inf \{B\} \leqq \inf \{A + B\}. \tag{8}$$

La comparaison des inégalités (7) et (8) fournit la relation (6) et un raisonnement analogue s'applique au cas des plus petites majorantes.

Notons que l'on aurait pu aussi déduire les théorèmes 2 et 3 de l'énoncé du théorème 1, sans faire appel aux détails de la démonstration.

Voici encore un corollaire du théorème 3.

Pour trois opérations linéaires positives A, B, C *on a toujours*

$$\inf (A, B + C) \leqq \inf (A, B) + \inf (A, C). \tag{9}$$

En effet, en remplaçant les ensembles $\{A\}$, $\{B\}$ par les couples A, B et A, C, la relation (6) s'écrit

$$\inf (2A, A + B, A + C, B + C) = \inf (A, B) + \inf (A, C)$$

et en observant que $A \leqq 2A$, $A \leqq A + B$ et $A \leqq A + C$ et que, par conséquent,

$$\inf (A, B + C) \leqq \inf (2A, A + B, A + C, B + C),$$

il en découle l'inégalité (9).

Disons encore quelques mots sur certains cas particuliers du théorème 1. Étant donné l'ensemble $\{A\}$, envisageons, pour chaque élément f, la borne inférieure des valeurs Af que nous désignerons par $\min \{Af\}$ pour la distinguer de $\inf \{A\}f$, valeur que prend l'opération $\inf \{A\}$ pour l'élément f. Ces deux opérations ne coïncident que si la première est distributive, c'est-à-dire que si $\min \{A(f + g)\} = \min \{Af\} + \min \{Ag\}$. Elles coïncident, entre autres, lorsque l'ensemble $\{A\}$ comprend, pour tout couple A_1, A_2 de ses éléments, au moins une minorante $A \leqq A_1$, $A \leqq A_2$. En effet, on voit immédiatement que sous cette hypothèse il y a aussi, dans l'ensemble $\{A\}$, une minorante A_0 pour tout n-uple $A_1, \cdots, A_n$ tiré de l'ensemble; par conséquent on a

$$\sum_1^m A_i f_i \geqq \sum_1^m A_0 f_i = A_0 f \geqq \min \{Af\},$$

et notre assertion s'en déduit immédiatement. On a donc le

THÉORÈME 4. *Lorsque l'ensemble $\{A\}$ comprend, pour chaque couple A_1, A_2 de ces éléments, un élément $A \leqq A_1$, $A \leqq A_2$, alors on a, pour tous les f,*

$$\inf \{A\}f = \min \{Af\}.$$

Évidemment, un fait analogue est valable pour les majorantes. Notons en particulier le cas où les éléments A dépendent d'un indice ou d'un paramètre continu λ allant vers l'infini et A_λ est monotone par rapport à λ, c'est-à-dire que l'on a, pour $\lambda < \mu$, ou bien constamment $A_\lambda \leqq A_\mu$ ou bien l'inégalité inverse. Dans ces cas, au lieu d'écrire $\inf \{A\}$ et $\sup \{A\}$, nous pourrons nous servir aussi de l'écriture évidente $\lim A_\lambda$, $A_\lambda \to \lim A_\lambda$ etc. Convenons encore d'écrire $B = \sum A_i$ dans le cas où les opérations positives A_i sont telles que la série $Bf = \sum A_i f$ est convergente pour tous les f; plus généralement, nous écrirons $B = \sum A$ aussi lorsque l'ensemble $\{A\}$ qui peut être infini non dénombrable, est tel que, pour chaque f, il n'y a qu'un nombre fini ou une infinité dénombrable d'opérations A pour lesquelles Af ne s'annule pas et que de plus la série formée de ces valeurs Af est convergente.

III. Opérations disjointes

Les opérations linéaires positives A et B seront dites d'être *disjointes* et nous écrirons $A \parallel B$ ou $B \parallel A$, lorsque $\inf (A, B) = 0$, c'est-à-dire que A et B n'admettent aucune minorante outre 0. D'après le théorème 2, cette hypothèse

s'écrit aussi $\sup(A, B) = A + B$. La construction de la plus grande minorante donnée au chapitre précédent met en évidence le critère suivant. *Pour que* $A \parallel B$, *il faut et il suffit que pour chaque élément* f *et pour chaque* $\epsilon > 0$ *il existe une décomposition* $f = f_1 + f_2$, *telle que*

$$Af_1 \leqq \epsilon, \qquad Bf_2 \leqq \epsilon.$$

Il suit immédiatement de la définition que si $A \parallel B$ et $A_1 \leqq A$, $B_1 \leqq B$, on a aussi $A_1 \parallel B_1$. Une autre conséquence immédiate consiste en ce que $A \parallel B$ et $a \geqq 0$, $b \geqq 0$ impliquent que $aA \parallel bB$. En effet, en posant $\max(a, b) = c$, on a

$$\inf(aA, bB) \leqq \inf(cA, cB) = c \inf(A, B) = 0.$$

THÉORÈME 5. *Lorsque* $A \parallel B$ *et* $A \leqq B + C$, *on a aussi* $A \leqq C$.

En effet, de $A \leqq B + C$ et $B \leqq B + C$ on déduit que

$$A + B = \sup(A, B) \leqq B + C,$$

donc $A \leqq C$.

THÉORÈME 6. *Lorsque* $A_k \parallel B$ $(k = 1, \cdots, n)$, *alors on a aussi* $\sum_1^n A_k \parallel B$.

Il suffit de considérer le cas $n = 2$, auquel le cas général se ramène par récurrence. Posons donc $C = \inf(A_1 + A_2, B)$; il nous faut montrer que $C = 0$. Or $C \leqq B \parallel A_1$, donc $C \parallel A_1$ et comme $C \leqq A_1 + A_2$, le théorème 5 fournit que $C \leqq A_2$. Par conséquent

$$C \leqq \inf(A_2, B) = 0.$$

THÉORÈME 7. *Lorsque l'ensemble* $\{A\}$ *est majorable et que l'on a* $A \parallel B$ *pour une opération* B *et pour tous les* A, *alors on a aussi* $\sup\{A\} \parallel B$.

Ce théorème est aussi un corollaire du théorème 5. Posons, en effet,

$$\inf(\sup\{A\}, B) = D;$$

alors on aura, pour tous les A,

$$A \leqq \sup\{A\} = D + (\sup\{A\} - D);$$

comme, d'autre part, $D \leqq B \parallel A$ entraîne que $A \parallel D$, on obtient du théorème 5 que

$$A \leqq \sup\{A\} - D$$

pour tous les A, ce qui donne $D = 0$.

Observons que *le théorème* 6 *s'étend d'une manière évidente au cas d'une infinité d'opérations* A_k, même au cas d'une infinité non dénombrable d'opérations A dont nous avons parlé à la fin du chapitre précédent; bien entendu on aura à supposer la convergence de $\sum A$. En effet, $\sum A$ n'est que la plus petite majorante de l'ensemble des opérations $\sum_1^m A_i$, formées de toutes les manières possibles d'éléments A_i tirés de l'ensemble $\{A\}$, et notre assertion se déduit de ce fait par les théorèmes 6 et 7. Un cas particulier important est celui où l'ensemble $\{A\}$, supposé majorable, est formé d'opérations A disjointes deux à

deux; dans ce cas on a évidemment $\sum A = \sup \{A\}$ et notre assertion est comprise dans le théorème 7.

THÉORÈME 8. *Pour deux opérations linéaires positives A, B, d'ailleurs quelconques, on a*

$$\sup (A, B) - A = B - \inf (A, B) \parallel A - \inf (A, B) = \sup (A, B) - B.$$

Pour le voir, posons $C = \inf (A_1, B_1)$ avec $A_1 = A - \inf (A, B)$, $B_1 = B - \inf (A, B)$; nous avons à montrer que $C = 0$. Or, comme $A_1 \geqq C$, $B_1 \geqq C$, nous obtenons

$$A = \inf (A, B) + A_1 \geqq \inf (A, B) + C,$$

$$B = \inf (A, B) + B_1 \geqq \inf (A, B) + C,$$

donc

$$\inf (A, B) \geqq \inf (A, B) + C$$

et de là $C = 0$.

THÉORÈME 9. *Lorsque $B \parallel C$, on a pour tout A*

$$\inf (A, B + C) = \inf (A, B) + \inf (A, C).$$

En effet, l'hypothèse faite assure que $\inf (A, B) \parallel \inf (A, C)$; de là et des relations évidentes

$$\inf (A, B + C) \geqq \inf (A, B), \qquad \inf (A, B + C) \geqq \inf (A, C)$$

on obtient

$$\inf (A, B + C) \geqq \sup (\inf (A, B), \qquad \inf (A, C)) = \inf (A, B) + \inf (A, C)$$

et la démonstration s'achève en comparant cette inégalité avec l'inégalité (9) du chapitre précédent.

IV. Parties disjointes des opérations linéaires

A et B étant des opérations linéaires positives telles que $A \leqq B$ et $A \parallel B - A$, nous convenons de dire que A est une *partie disjointe* de B et d'écrire $A \prec B$. Il est manifeste qu'on aura en même temps $B - A \prec B$.

THÉORÈME 10. *Lorsque $A \prec B \prec C$, on a aussi $A \prec C$.*

En effet, comme $B \parallel C - B$ et $A \leqq B$, on aura aussi $A \parallel C - B$; comme de plus $A \parallel B - A$, on déduit du théorème 6 que

$$A \parallel (C - B) + (B - A) = C - A$$

c'est-à-dire que $A \prec C$.

THÉORÈME 11. *Lorsque $A \leqq B \leqq C$ et $A \prec C$, on a aussi $A \prec B$.*

En effet, d'après les hypothèses faites, on a $A \parallel C - A$ et $B - A \leqq C - A$ et par conséquent $A \parallel B - A$, c'est-à-dire que $A \prec B$.

THÉORÈME 12. *Lorsque $A \prec B$ pour tous les éléments A de l'ensemble $\{A\}$, alors* $\inf \{A\} \prec B$ *et* $\sup \{A\} \prec B$.

Posons inf $\{A\} = C$, sup $\{A\} = D$. Alors par l'hypothèse faite $C \leqq A \parallel B - A$ pour tout A, donc $C \parallel B - A$ et par le théorème 7 il s'ensuit que

$$C \parallel \sup \{B - A\} = B - \inf \{A\} = B - C,$$

c'est-à-dire que inf $\{A\} = C \prec B$. Pour obtenir de là la seconde assertion du théorème, on n'a qu'à appliquer le résultat obtenu à l'ensemble $\{B - A\}$; en effet,

$$B - D = B - \sup \{A\} = \inf \{B - A\} \prec B,$$

donc aussi $D \prec B$.

THÉORÈME 13. *Lorsque $A \not\equiv 0$, $A \prec B$ et que a et b sont des coefficients numériques non-négatifs tels que $aA \leqq bB$, alors on a aussi $a \leqq b$.*

En effet, de $A \parallel B - A$ il s'ensuit que $aA \parallel b(B - A)$; comme de plus $b(B - A) \leqq bB$ et $aA \leqq bB$, on obtient que

$$aA + b(B - A) = \sup (aA, b(B - A)) \leqq bB$$

c'est-à-dire que $aA \leqq bA$. Or, soit f un élément tel que $Af \not\equiv 0$, donc $Af > 0$, alors $aAf \leqq bAf$ donne que $a \leqq b$, C. Q. F. D.

V. Familles complètes et familles presque complètes

Un ensemble d'opérations linéaires positives sera dit de former une *famille complète* lorsqu'il comprend

1) toutes les minorantes de chacun de ses éléments,
2) la somme de deux quelconques de ses éléments et
3) les plus petites majorantes de ses sous-ensembles majorables.

Voici quelques exemples: a) l'ensemble de toutes les opérations linéaires positives d'un domaine fondamental Ω; b) l'ensemble de toutes les opérations qui s'annulent pour certains éléments f donnés, en nombre fini ou infini; c) l'ensemble des opérations qui sont disjointes à certaines opérations données, en nombre fini ou infini. L'exemple a) est évident et les hypothèses 1) et 2) sont aussi évidemment réalisées par l'exemple b); pour vérifier dans ce cas l'hypothèse 3), nous n'avons qu'à rappeler le procédé par lequel on construit la plus petite majorante. Quant à l'exemple c), l'hypothèse 1) est une conséquence immédiate de la définition des opérations disjointes et les hypothèses 2) et 3) se vérifient par les théorèmes 6 et 7.

THÉORÈME 14. *Lorsque l'ensemble $\{A\}$ forme une famille complète, alors toute opération linéaire positive B se décompose, d'une façon unique, en deux parties disjointes A_1 et $B_1 = B - A_1$ où A_1 appartient à la famille $\{A\}$ et $B_1 \parallel A$ pour tous les A. Les opérations A_1, B_1 ne sont que les plus grandes parties disjointes de l'opération B, la première parmi celles qui appartiennent à la famille $\{A\}$, la seconde parmi celles qui sont disjointes à tous les A.*

Pour démontrer ce théorème, désignons par A_1 la plus petite majorante de ceux des éléments A de l'ensemble $\{A\}$ pour lesquels $A \leqq B$, et posons $B_1 = B - A_1$. Alors, par l'hypothèse 3), A_1 appartient à l'ensemble $\{A\}$ et il nous

faut encore montrer que $B_1 \,||\, A$ pour tous les A. Or, par les hypothèses 1) et 2), les opérations

$$A_1 + \inf (A, B_1)$$

appartiennent aussi à l'ensemble $\{A\}$; comme de plus

$$A_1 + \inf (A, B_1) \leqq A_1 + B_1 = B$$

alors, de par la définition de A_1,

$$A_1 + \inf (A, B_1) \leqq A_1$$

et par conséquent, $\inf (A, B_1) = 0$, c'est-à-dire que $B_1 \,||\, A$, ce qu'il fallait démontrer.

L'unicité s'ensuit immédiatement du théorème 5.

La dernière assertion du théorème vient immédiatement de notre démonstration en ce qu'elle concerne l'opération A_1. Quant à l'opération B_1, soit B_2 une partie disjointe de B (ou plus généralement soit $B_2 \leqq B$) et supposons que $B_2 \,||\, A$ pour tous les A. Alors comme $B_2 \leqq B = A_1 + B_1$ et $B_2 \,||\, A_1$, il suit du théorème 5 que $B_2 \leqq B_1$ et comme encore B_1 est elle-même une partie disjointe de B, notre assertion est démontrée.

La décomposition d'une opération linéaire portant sur les fonctions continues en des parties appartenant à trois classes particulières importantes dont j'ai parlé dans ma conférence de Bologne, est une des applications les plus caractéristiques du théorème 14.

THÉORÈME 15. *Étant donné l'ensemble $\{A\}$, désignons par A' les opérations qui sont disjointes à toutes les opérations A et par A'' celles disjointes à toutes les A'. Alors les ensembles $\{A'\}$ et $\{A''\}$ forment des familles complètes et $\{A''\}$ est la plus petite famille complète embrassant l'ensemble $\{A\}$.*

Observons tout d'abord que l'existence d'une plus petite parmi les familles complètes embrassant l'ensemble $\{A\}$ est évidente, cette plus petite famille n'étant que la partie commune de toutes ces familles complètes.

Quant à l'ensemble $\{A'\}$, pour voir qu'il forme une famille complète, nous n'avons qu'à rappeler notre exemple c); il en est de même pour l'ensemble $\{A''\}$. Quant à la dernière assertion du théorème, soit $[A]$ la plus petite famille complète embrassant l'ensemble $\{A\}$; il faut montrer que $\{A''\} \equiv [A]$. D'abord il est manifeste que la famille $\{A''\}$ comprend l'ensemble $\{A\}$; ce qui reste à prouver c'est que toute opération A'' est comprise dans la famille $[A]$. A cet effet, soit A_1'' l'une quelconque des opérations A'' et mettons-la, conformément au théorème 14, sous la forme $A_1'' = A_1 + B_1$, A_1 appartenant à la famille $[A]$ et B_1 étant disjoint à tous les éléments de cette famille, donc à plus forte raison à tous les éléments de l'ensemble $\{A\}$, c'est-à-dire que B_1 est un A'. Comme encore $B_1 \leqq A_1''$, alors B_1 appartient aussi à la famille $\{A''\}$. Par conséquent $B_1 \,||\, B_1$, donc $B_1 = 0$, c'est-à-dire que $A_1'' = A_1$ et ainsi A_1'' appartient à la famille $[A]$, ce qu'il fallait démontrer.

Observons encore que lorsque l'ensemble $\{A\}$ forme lui-même une famille

complète et que par conséquent $\{A\} \equiv [A]$, notre théorème se réduit au corollaire suivant. *A toute famille complète* $\{A\}$ *il correspond une famille complémentaire* $\{A'\}$ *bien déterminée, formée par les opérations disjointes à toutes les opérations* A, *et la famille complementaire qui correspond à la famille* $\{A'\}$ *est fournie par la famille primitive* $\{A\}$ *elle-même.*

Voici encore un second corollaire, énoncé sous une forme particulière dont nous aurons besoin dans la suite. *Lorsque* $A \parallel B$ *et que* $[A]$ *et* $[B]$ *désignent les plus petites familles complètes embrassant respectivement les opérations* A *et* B, *alors* $[A] \parallel [B]$, *ce qui veut dire que tout élément de la famille* $[A]$ *est disjoint à tout élément de la famille* $[B]$.

Ici peut-être convient-il d'observer que la plus petite famille complète $[A]$ déterminée par une seule opération donnée A peut être très peuplée. Voici un des exemples les plus caractéristiques, cas particulier de l'exemple 7 dont nous avons parlé au chapitre I. Soit Ω l'ensemble des suites $x = (x_k)$ formées de termes non-négatifs, à somme finie, l'addition de deux suites se faisant comme d'ordinaire. On voit aisément que les opérations linéaires positives A dans Ω s'expriment sous la forme $Ax = \sum a_k x_k$ où (a_k) est une suite bornée à termes positifs, déterminée par l'opération A et que, inversement, toute telle suite (a_k) donne naissance à une opération A. Aussi est-il évident que la plus petite minorante des opérations $A = (a_k)$ et $B = (b_k)$ correspond à la suite $(\min (a_k, b_k))$ et que, par conséquent, l'opération $A_1 = \sum x_k$ pour laquelle $a_1 = \cdots = 1$ n'est disjointe à aucune des opérations sauf 0. C'est-à-dire que dans le cas envisagé la plus petite famille complète $[A_1]$ qui embrasse A_1, épuise toutes les opérations possibles. Observons encore que pour le domaine fondamental Ω figurant dans l'exemple 1, formé par les fonctions continues non-négatives $f(x)$ définies dans l'intervalle $a \leqq x \leqq b$, la plus petite famille complète $[I]$ embrassant l'intégration

$$If = \int_a^b f(x)\,dx \tag{1}$$

n'est que l'ensemble des opérations

$$\int_a^b g(x) f(x)\,dx \tag{2}$$

correspondant à des fonctions $g(x)$ non-négatives et intégrables au sens de Lebesgue, d'ailleurs quelconques.

Disons maintenant quelques mots des *familles presque complètes* que nous définissons en supprimant l'hypothèse 3). L'ensemble $\{A\}$ sera donc dit de former une famille presque complète lorsqu'il comprend

1) toutes les minorantes de chacun de ses éléments et

2) la somme de deux quelconques de ses éléments.

Une famille presque complète, de même que les familles complètes, contient la plus petite majorante d'un nombre fini de ses éléments, $A_1, \cdots, A_n$, ce qui suit de la définition en observant que cette majorante est une minorante de la

somme $A_1 + \cdots + A_n$, mais ce fait ne subsiste plus nécéssairement pour une infinité d'éléments.

L'existence d'une plus petite famille presque complète embrassant l'ensemble donné $\{A\}$ est manifeste et cela pour les mêmes raisons que pour la famille complète. De plus l'hypothèse 2) met en évidence que cette famille devra comprendre toutes les opérations de la forme

$$a_1A_1 + \cdots + a_nA_n \tag{3}$$

où $A_1, \cdots, A_n$ sont des opérations appartenant à l'ensemble $\{A\}$ et $a_1, \cdots, a_n$ désignent des nombres entiers positifs. D'autre part, d'après l'hypothèse 1), notre famille devra aussi comprendre toutes les minorantes des opérations (3). Mais il est évident que l'ensemble de ces minorantes satisfait aux hypothèses 1) et 2), c'est-à-dire qu'il fournit précisément la plus petite famille en question.

En particulier, la plus petite famille presque complète embrassant une opération donnée A se compose de toutes les minorantes de A et de leurs multiples ou ce qui revient au même, elle n'est que l'ensemble des opérations B bornées par rapport à A, $B \leqq cA$. Ainsi par exemple, pour l'opération I donnée par la formule (1), c'est-à-dire l'intégration, la plus petite famille presque complète qui l'embrasse est formée par celles des opérations (2) qui correspondent à des fonctions $g(x)$ bornées.

Convenons encore d'appeler, dans la suite, *famille de l'opération* A la plus petite famille complète qui l'embrasse. Nous venons de voir que cette famille, et même la plus petite famille presque complète embrassant une opération, peuvent engendrer toutes les opérations du domaine envisagé. Parlons maintenant de l'autre possibilité extrême. Supposons que l'opération A n'admette pas d'autres minorantes que celles de la forme cA $(0 \leqq c \leqq 1)$, alors il est manifeste que les deux familles en question se bornent aux mêmes opérations cA; bien entendu, c varie de 0 à l'infini. Tel est le cas dans nos exemples 1)–3) pour l'opération A définie par $Af = cf(a)$ ou plus généralement par $Af = cf(x_0)$ $(c \geqq 0)$. Quant à l'exemple 4), celui-ci ne donne pas lieu à de telles opérations "singulières."

Enfin observons que l'on peut passer, comme on montre immédiatement, de la famille presque complète déterminée par une opération A à sa famille $[A]$, et cela en y ajoutant les plus petites majorantes des sous-ensembles majorables. Il en est de même pour les plus petites familles complète et presque complète qui embrassent un ensemble $\{A\}$ d'opérations données.

VI. Opérations linéaires de signe quelconque

Passons aux opérations de la forme $A = B - C$ où B et C sont des opérations linéaires positives mais sans que l'on ait nécessairement $C \leqq B$. Alors l'opération A ne sera pas nécessairement positive; convenons de l'appeler tout simplement opération linéaire.

Pour qu'une opération A appartienne à la classe envisagée, il faut et il suffit qu'elle soit distributive, $A(f + g) = Af + Ag$, et que de plus elle soit majorable

par une opération linéaire positive, c'est-à-dire qu'il existe une opération linéaire positive B de sorte que l'on ait $Af \leqq Bf$ pour tous les f. La condition est évidente.

L'opération $A = B - C$ où $B \geqq 0$, $C \geqq 0$ peut s'écrire sous la forme canonique $A = B_1 - C_1$, en posant

$$(1) \qquad \begin{aligned} B_1 &= B - \inf(B, C) = \sup(B, C) - C, \\ C_1 &= C - \inf(B, C) = \sup(B, C) - B; \end{aligned}$$

les opérations B_1, C_1 s'appelleront respectivement les *parties positive et négative de* A et nous écrirons aussi pos A au lieu de B_1 et neg A au lieu de C_1. Je dis que ces opérations B_1 et C_1 sont disjointes et qu'il n'existe pas d'autre décomposition $A = B_2 - C_2$, $B_2 \geqq 0$, $C_2 \geqq 0$ avec B_2, C_2 disjoints. En effet, la relation $B_1 \parallel C_1$ est assurée par le théorème 8, et si l'on avait encore $B_2 \parallel C_2$, alors comme

$$(2) \qquad B_1 + C_2 = B_2 + C_1$$

et par conséquent $B_1 \leqq B_2 + C_1$, il suivrait du théorème 5 que $B_1 \leqq B_2$ et par symétrie que $B_2 \leqq B_1$, donc $B_1 = B_2$ et alors, par (2), $C_1 = C_2$.

Il s'ensuit que les opérations $B_1 = \text{pos } A$, $C_1 = \text{neg } A$ sont déterminées par l'opération A elle-même, indépendamment de la décomposition $A = B - C$ dont nous sommes partis. De là il vient encore par (1) que pos $A \leqq B$, neg $A \leqq C$ pour toute telle décomposition et comme $A = \text{pos } A - \text{neg } A$ est elle-même une telle décomposition, alors

$$\text{pos } A = \inf\{B\}, \qquad \text{neg } A = \inf\{C\},$$

formés avec les B et C qui figurent dans l'ensemble des décompositions possibles.

Les notions de minorante et de majorante, de plus grande minorante et de plus petite majorante s'étendent d'une manière évidente aux opérations linéaires de signe quelconque; cela étant, on pourra aussi écrire

$$\text{pos } A = \sup(A, 0); \qquad \text{neg } A = \sup(-A, 0).$$

En effet, toute majorante B des opérations A et 0 ou ce qui revient au même, toute majorante positive B de l'opération A, fournit une décomposition $A = B - C$, $B \geqq 0$, $C = B - A \geqq 0$ et de même, toute majorante positive C de $-A$ donne lieu à une telle décomposition avec $B = A + C = C - (-A) \geqq 0$; de là notre assertion s'ensuit immédiatement.

En résumé, on a le théorème suivant.

THÉORÈME 16. *Toute opération linéaire A de signe quelconque peut être écrite, d'une seule manière, comme différence de deux opérations linéaires positives et disjointes l'une à l'autre, savoir*

$$A = \text{pos } A - \text{neg } A; \qquad \text{pos } A = \sup(A, 0), \qquad \text{neg } A = \sup(-A, 0).$$

Pour toute autre décomposition $A = B - C$, $B \geqq 0$, $C \geqq 0$ on a

$$B \geqq \text{pos } A, \qquad C \geqq \text{neg } A.$$

En ce qui concerne la plus grande minorante et la plus petite majorante d'un ensemble $\{A\}$, on peut les construire par le même procédé que dans le cas des opérations positives et il est aussi manifeste que les théorèmes 1–4 restent valables, avec la seule modification que les ensembles à une infinité d'éléments ne sont pas nécessairement minorables.

On pourra aussi étendre au cas envisagé la notion des opérations disjointes et cela en convenant de considérer comme disjointes les opérations A et B et d'écrire $A \parallel B$ lorsque l'on a simultanément

$$\text{pos } A \parallel \text{pos } B, \qquad \text{pos } A \parallel \text{neg } B, \qquad \text{neg } A \parallel \text{pos } B, \qquad \text{neg } A \parallel \text{neg } B$$

ou plus brièvement, avec l'écriture

$$\text{pos } A + \text{neg } A = \text{mod } A,$$

toujours que

$$\text{mod } A \parallel \text{mod } B.$$

On voit aisément que les théorèmes 6–8 concernant les opérations disjointes restent valables sous ces conditions plus générales. En effet, quant au théorème 6, le passage du cas des opérations positives au cas général se fait immédiatement à l'aide des relations

$$\text{pos (sup } \{A\}) = \text{sup } \{\text{pos } A\}; \qquad \text{neg (inf } \{A\}) = \text{sup } \{\text{neg } A\}$$

qui à leur tour sont des conséquences des inégalités évidentes

$$\text{pos } A \leqq \text{pos (sup } \{A\}); \qquad \text{sup } \{A\} \leqq \text{sup } \{\text{pos } A\} \geqq 0$$

et de celles qui en ressortent lorsqu'on passe à l'ensemble $\{-A\}$. L'extension du théorème 7 se déduit des inégalités évidentes

$$(2) \qquad \text{pos} \sum_1^n A_k \leqq \sum_1^n \text{pos } A_k; \qquad \text{neg} \sum_1^n A_k \leqq \sum_1^n \text{neg } A_k.$$

Enfin, pour étendre le théorème 8, on n'a qu'à l'appliquer aux opérations pos A + neg B et neg A + pos B au lieu de A et B et puis retrancher neg A + neg B de chaque terme de la formule qu'on obtient.

Observons en passant que pour des opérations A_k disjointes entre elles, c'est toujours le signe d'égalité qui est valable dans les formules (2). Cela vient en écrivant $\sum A_k = \sum \text{pos } A_k - \sum \text{neg } A_k$ et en observant que, par l'hypothèse faite et par le théorème 7 les deux termes du second membre représentent des opérations linéaires positives et disjointes.

Pour étendre la notion de partie disjointe, convenons de dire que A est une partie disjointe de B et d'écrire $A \prec B$ lorsque $A \parallel B - A$, tout comme pour des opérations positives. On voit aisément que, avec la convention posée, les théorèmes 10 et 12 subsistent.

Quant aux familles complètes et presque complètes, réservons ces dénominations pour les familles d'opérations positives et appelons prolongement de la

famille ce qui en vient en y ajoutant les différences de ses éléments. Il est manifeste qu'on obtient déjà toutes ces différences, chacune en un seul exemplaire, quand on se contente de former les différences à termes disjoints. On voit aussi immédiatement qu'une famille et son prolongement embrassent, avec A et B, leurs combinaisons linéaires $aA + bB$ où a et b sont des facteurs numériques de signe quelconque, ainsi que les opérations inf (A, B) et sup (A, B). Ce dernier fait résulte par exemple des formules

$$\inf(A,B) = \inf(A + C, B + C) - C, \quad \sup(A,B) = \sup(A + C, B + C) - C$$

en posant $C = \text{neg } A + \text{neg } B$ et en rappelant que le fait analogue, pour les opérations positives $A + C$, $B + C$, est impliqué dans la définition des familles complètes ou presque complètes.

VII. Décomposition simultanée des opérations linéaires

Le premier pas vers un calcul fonctionnel des opérations linéaires est la décomposition simultanée de deux opérations linéaires A et E dont la seconde est positive; cette opération E et certaines de ses parties disjointes auront à jouer dans la suite le rôle d'une sorte d'unités. Nous supposons encore que l'opération A appartienne à la famille $[E]$ de l'opération E ou bien, lorsque l'opération A n'est pas positive, au prolongement de cette famille ou ce qui revient au même que toute opération disjointe à E le soit aussi par rapport à A.

Soit λ une variable réelle et envisageons les opérations

$$P_\lambda = \text{pos}(\lambda E - A),$$

$$N_\lambda = \text{neg}(\lambda E - A) = P_\lambda - \lambda E + A$$

en fonction de λ. Elles sont positives et monotones, la première croissant avec λ et la seconde décroissant, c'est-à-dire que, pour $\lambda_1 < \lambda_2$, on a $P_{\lambda_1} \leqq P_{\lambda_2}$ et $N_{\lambda_1} \geqq N_{\lambda_2}$. Je dis que P_λ et N_λ sont des fonctions convexes; cela veut dire que les opérations

$$\frac{P_{\lambda+h} - P_\lambda}{h}, \qquad \frac{N_{\lambda+h} - N_\lambda}{h}$$

sont des fonctions croissantes de λ. Notre assertion est évidemment équivalente à celle que pour tout élément f, $P_\lambda f$ et $N_\lambda f$ sont des fonctions numériques convexes. Or, pour démontrer cette dernière assertion, nous n'avons qu'à observer que la fonction $P_\lambda f$ est précisément l'enveloppe supérieure des parties positives des fonctions numériques linéaires $\lambda Eg - Ag$, formées avec tous les $g \leqq f$, c'est-à-dire pour toutes les décompositions $f = g + (f - g)$. Or toutes ces parties positives sont des fonctions convexes et par conséquent il en est de même pour leur enveloppe supérieure. La convexité des fonctions $N_\lambda f$ s'en déduit par la formule $N_\lambda = P_\lambda - \lambda E + A$.

Les fonctions $P_\lambda f$ et $N_\lambda f$ étant convexes, elles admettent partout des dérivées des deux côtés. Envisageons seulement les dérivées à droite; comme fonctions de λ, elles sont croissantes et continues à droite. Comme fonctions de f, elles

définissent, pour chaque valeur de λ, deux opérations linéaires que nous désignerons par P'_λ et N'_λ. Celles-ci sont évidemment des fonctions croissantes de λ et comme P_λ est croissante et N_λ décroissante, alors $P'_\lambda \geqq 0$, $N'_\lambda \leqq 0$.

Les opérations P_λ et N_λ sont disjointes de par le théorème 16, $P_\lambda \parallel N_\lambda$. Nous allons montrer que l'on a aussi $P_\lambda \parallel N'_\lambda$, $P'_\lambda \parallel N_\lambda$, $P'_\lambda \parallel N'_\lambda$. D'abord, N_λ étant positive et décroissante, on a

$$0 \leqq N_\lambda - N_{\lambda+h} \leqq N_\lambda \qquad (h > 0),$$

donc $P_\lambda \parallel N_\lambda$ entraîne que $P_\lambda \parallel N_\lambda - N_{\lambda+h}$ et par conséquent

$$P_\lambda \parallel \frac{N_\lambda - N_{\lambda+h}}{h};$$

de là il suit que $P_\lambda \parallel -N'_\lambda$ ou alors $P_\lambda \parallel N'_\lambda$, le passage à la limite étant assuré par le théorème 7. Soit de plus $\lambda < \mu$, donc $\lambda + h \leqq \mu$ pour h suffisemment petit; alors comme $P_{\lambda+h} \parallel N_{\lambda+h} \geqq N_\mu \geqq 0$ et comme P_λ est positive, croissante et convexe, il vient que

$$P'_\lambda \leqq \frac{P_{\lambda+h} - P_\lambda}{h} \parallel N_\mu,$$

ce qui donne que $P'_\lambda \parallel N_\mu$; de là on passe à $P'_\lambda \parallel N_\lambda$ en se servant du théorème 7. Enfin un raisonnement analogue fournit

$$P'_\lambda \parallel \frac{N_\lambda - N_{\lambda+h}}{h}$$

donc aussi $P'_\lambda \parallel - N'_\lambda$ et alors $P'_\lambda \parallel N'_\lambda$.

Cela étant, différentions la relation $P_\lambda - N_\lambda = \lambda E - A$; on obtient que $P'_\lambda - N'_\lambda = E$, donc on pourra poser, par définition,

$$E_\lambda = P'_\lambda, \qquad E - E_\lambda = -N'_\lambda \tag{1}$$

et *les opérations E_λ et $E - E_\lambda$ sont des parties disjointes de E. La décomposition correspondante de l'opération A en parties disjointes est fournie par*

$$A_\lambda = \lambda P'_\lambda - P_\lambda, \qquad A - A_\lambda = -\lambda N'_\lambda + N_\lambda. \tag{2}$$

En effet, la somme des seconds membres donne A; on le vérifie en rappelant les relations $P_\lambda - N_\lambda = \lambda E - A$, $P'_\lambda - N'_\lambda = E$. Le fait que $A_\lambda \parallel A - A_\lambda$ s'obtient par le théorème 6 des relations entre P_λ, N_λ et leurs dérivées que nous venons de démontrer. Par les mêmes raisons *on a aussi* $E_\lambda \parallel A - A_\lambda$ *et* $A_\lambda \parallel E - E_\lambda$.

Établissons encore une inégalité relative aux opérations E_λ, A_λ et importante pour la suite. L'opération P_λ étant une fonction convexe de λ, on a, pour $\lambda < \mu$,

$$(\mu - \lambda)P'_\mu \geqq P_\mu - P_\lambda \geqq (\mu - \lambda)P'_\lambda.$$

En ajoutant $\lambda P'_\lambda - \mu P'_\mu$ à chaque membre et en changeant les signes, nous obtenons

$$\lambda(P'_\mu - P'_\lambda) \leqq \mu P'_\mu - P_\mu - \lambda P'_\lambda + P_\lambda \leqq \mu(P'_\mu - P'_\lambda)$$

ou avec les notations (1) et (2)

$$\lambda(E_\mu - E_\lambda) \leqq A_\mu - A_\lambda \leqq \mu(E_\mu - E_\lambda). \tag{3}$$

Calculons encore les limites de E_λ et A_λ pour λ infini, positif et négatif. Pour $0 < \lambda \to \infty$ on a

$$E \geqq E_\lambda = P'_\lambda \geqq (P_\lambda - P_0)/\lambda \geqq (\lambda E - A - P_0)/\lambda \to E,$$

ce qui donne $E_\infty = E$. De même, pour $0 > \lambda \to -\infty$, on a

$$E \geqq E - E_\lambda = -N'_\lambda \geqq -(N_\lambda - N_0)/\lambda \geqq (\lambda E - A + N_0)/\lambda \to E$$

et par conséquent $E_{-\infty} = 0$. De plus comme $N_\lambda \geqq 0$ et $N'_\lambda \leqq 0$, l'opération $A - A_\lambda = -\lambda N'_\lambda + N_\lambda$ est positive pour $\lambda > 0$ et sous la même hypothèse, d'après l'inégalité (3), A_λ croît avec λ. Donc pour $\lambda \to \infty$, A_λ tend vers une opération limite A_∞. Comme de plus $E_\lambda \,||\, A - A_\lambda \geqq A - A_\infty$, donc $E_\lambda \,||\, A - A_\infty$ et par le théorème 7 il en ressort que $E = E_\infty = \sup E_\lambda \,||\, A - A_\infty$. Un raisonnement analogue donne l'existence de $A_{-\infty}$ et la relation $E \,||\, A_{-\infty}$. Or, par hypothèse, A appartient à la famille $[E]$ de E ou à son prolongement et on voit successivement qu'il en est de même pour P_λ, N_λ, P'_λ, N'_λ, donc aussi pour A_λ et enfin pour ses limites A_∞ et $A_{-\infty}$. Ce fait et les relations $E \,||\, A - A_\infty$, $E \,||\, A_{-\infty}$ que nous venons de prouver, ne sont compatibles que si $A - A_\infty = 0$, $A_{-\infty} = 0$.

En résumé, on a $E_\infty = E$, $A_\infty = A$, $E_{-\infty} = A_{-\infty} = 0$.

Enfin considérons les opérations

$$E_0 = P'_0, \qquad E - E_0 = -N'_0, \qquad A_0 = -P_0 = -\sup(-A, 0) = -\operatorname{neg} A,$$

$$A - A_0 = \operatorname{pos} A.$$

L'opération E_0 ou plutôt la décomposition $E = E_0 + (E - E_0)$ se distingue, parmi les décompositions de E en parties disjointes, par chacun des deux faits suivants. E_0 est la plus grande partie disjointe de E qui est disjointe à l'opération pos A et $E - E_0$ est la plus grande partie de E appartenant à la famille de pos A. Comme nous savons déjà que $E_0 \,||\, A - A_0 = \operatorname{pos} A$, nos assertions découlent du théorème 14, après avoir montré que l'opération $E - E_0$ appartient à la famille de pos A, c'est-à-dire qu'elle est disjointe à toute opération B disjointe à pos A. Or, pour une telle opération B, comme, par l'inégalité (3), pos $A = A - A_0 \geqq A_\mu - A_\lambda \geqq \lambda(E_\mu - E_\lambda)$ pour $0 < \lambda < \mu$, il vient que $B \,||\, E_\mu - E_\lambda$ et pour $\lambda \to 0$, $\mu \to \infty$ il en découle, par le théorème 7, que $B \,||\, E_\infty - E_0 = E - E_0$ ce qu'il fallait démontrer.

Du cas $\lambda = 0$ que nous venons d'envisager, on passe au cas général en remplaçant λ par $\lambda - \lambda_0$ et A par $A - \lambda_0 E$ et l'on obtient la caractérisation suivante

de la décomposition $E = E_{\lambda_0} + (E - E_{\lambda_0})$. *Les opérations* E_{λ_0} *et* $E - E_{\lambda_0}$ *sont les plus grandes parties disjointes de* E, *la première parmi celles qui sont disjointes à* pos $(A - \lambda_0 E)$ *et la seconde parmi celles comprises dans la famille de cette dernière.* Il vient en particulier que $E_\lambda = 0$ pour les valeurs négatives de λ toujours quand l'opération A est positive et plus généralement que $E_\lambda = 0$ pour $\lambda < a$ lorsque $A \geqq aE$. Un autre cas particulier intéressant se présente quand A est une partie disjointe de E; un calcul facile donne que dans ce cas $E_\lambda = 0, E - A, E$ suivant que $\lambda < 0, 0 \leqq \lambda < 1, \lambda \geqq 1$ et $A_\lambda = 0$ pour $\lambda < 1$ et $A_\lambda = A$ pour $\lambda \geqq 1$.

Nous aurons encore besoin dans la suite d'une *décomposition à deux paramètres* de E, induite par deux opérations A et B. Soit B une opération linéaire, appartenant, tout comme A, à la famille de l'opération E ou à son prolongement et faisons la décomposition simultanée de E et B comme nous venons de la faire pour E et A. Pour distinguer, nous désignons par μ le paramètre qui intervient et par $\bar{E}_\mu$ la décomposition correspondante de E. Envisageons l'opération

$$E_{\lambda,\mu} = \inf (E_\lambda , \bar{E}_\mu) \tag{4}$$

comme fonction des deux variables λ et μ et indiquons la propriété caractéristique qui distingue l'opération $E_{\lambda,\mu}$ parmi les parties disjointes de l'opération E. Elle suit immédiatement de nos résultats. L'opération $E_{\lambda,\mu}$ est la plus grande parmi celles des parties disjointes de E qui sont disjointes à la fois aux opérations $C = \text{pos}\,(A - \lambda E)$ et $D = \text{pos}\,(B - \mu E)$ et par conséquent, $E - E_{\lambda,\mu}$ est la plus grande partie disjointe de E comprise dans la plus petite famille complète embrassant les opérations C et D.

Notons encore que l'on a, pour $-\infty \leqq \lambda_1 < \lambda_2 \leqq \infty$, $-\infty \leqq \mu_1 < \mu_2 \leqq \infty$, la relation

$$E_{\lambda_1,\mu_1} + E_{\lambda_2,\mu_2} - E_{\lambda_1,\mu_2} - E_{\lambda_2,\mu_1} = \inf (E_{\lambda_2} - E_{\lambda_1} , \bar{E}_{\mu_2} - \bar{E}_{\mu_1}), \tag{5}$$

que l'on déduit du théorème 9 en passant d'abord, sous l'hypothèse $A_1 \,||\, A_2$, $B_1 \,||\, B_2$, à la formule

$$\inf (A_1 + A_2 , B_1 + B_2) = \inf (A_1 , B_1 + B_2) + \inf (A_1 + A_2 , B_1) + \inf (A_2 , B_2) - \inf (A_1 , B_1)$$

et puis en y remplaçant A_1, A_2, B_1 et B_2 respectivement par E_{λ_1}, $E_{\lambda_2} - E_{\lambda_1}$, $\bar{E}_{\mu_1}$ et $\bar{E}_{\mu_2} - \bar{E}_{\mu_1}$.

VIII. Calcul fonctionnel des opérations linéaires

Envisageons comme au chapitre précédent une opération linéaire positive E et une opération linéaire A, de signe quelconque, appartenant à la famille de E ou à son prolongement et soit E_λ la décomposition de E que nous venons de considérer. Supposons que la fonction $\Phi(\lambda)$ soit intégrable par rapport à E_λ, ce qui veut dire que la fonction Φ est intégrable, au sens de Stieltjes-Lebesgue, pour tous les f, par rapport aux fonctions $E_\lambda f$, fonctions numériques non-dé-

croissantes et continues à droite de la variable λ.[6] Lorsque la fonction Φ est non-négative, il est manifeste qu'en faisant correspondre à chaque élément f la valeur fournie par l'intégrale respective, on aura défini une opération linéaire positive que l'on désignera par $\Phi_E(A)$, pour mettre ainsi en évidence la dépendance de la nouvelle opération de E, A et de la fonction $\Phi(\lambda)$; on écrira aussi, avec une notation évidente,

$$\Phi = \Phi_E(A) = \int_{-\infty}^{\infty} \Phi(\lambda)\, dE_\lambda. \tag{1}$$

Lorsque la fonction Φ est de signe quelconque, l'opération définie de la même manière est encore une opération linéaire, étant évidemment égale à la différence des deux opérations positives qui correspondent aux parties positive et négative de la fonction Φ.

Une clef à l'analyse fonctionnelle des opérations du type $\Phi_E(A)$ est fournie par le théorème suivant.

THÉORÈME 17. *Les fonctions non-négatives* $\Phi_1(\lambda)$ *et* $\Phi_2(\lambda)$ *étant intégrables par rapport à* E_λ *et de plus telles que* $\Phi_1(\lambda)\Phi_2(\lambda) \equiv 0$, *les opérations* $\Phi_{1E}(A)$ *et* $\Phi_{2E}(A)$ *seront disjointes l'une à l'autre.*

Pour le démontrer, commençons par le cas particulier où Φ_1 et Φ_2 sont les fonctions caractéristiques de deux ensembles e_1 et e_2 sans point commun et supposons d'abord que l'ensemble e_1 soit la réunion d'un nombre fini ou d'une infinité dénombrable d'intervalles demiouverts $\lambda_i < \lambda \leqq \lambda_i'$ n'empiétant pas. Alors on a

$$\Phi_1 = \Phi_{1E}(A) = \sum_i (E_{\lambda_i'} - E_{\lambda_i}) \tag{2}$$

et

$$\Phi_2 = \Phi_{2E}(A) \leqq E - (E_{\lambda_i'} - E_{\lambda_i}) \qquad (i = 1, 2, \cdots). \tag{3}$$

Or de $E_{\lambda_i} \prec E$, $E_{\lambda_i'} \prec E$ et $E_{\lambda_i} \leqq E_{\lambda_i'} \leqq E$ il vient immédiatement, par les théorèmes 11 et 10, que $E_{\lambda_i'} - E_{\lambda_i} \prec E$, c'est-à-dire que $E_{\lambda_i'} - E_{\lambda_i} \,||\, E - (E_{\lambda_i'} - E_{\lambda_i})$. Donc par la formule (3) on a aussi $E_{\lambda_i'} - E_{\lambda_i} \,||\, \Phi_2$ et de là il s'ensuit, par la formule (2) et par les théorèmes 6 et 7, que $\Phi_1 \,||\, \Phi_2$, ce qu'il fallait démontrer.

Passons au cas où l'ensemble e_1 n'est plus supposé d'être composé d'intervalles; alors en fixant l'élément f et en choisissant un $\epsilon > 0$, on pourra, comme on sait bien, enfermer e_1 dans un ensemble e_3 du type particulier que nous venons de considérer et cela de sorte que les mesures des deux ensembles, par rapport à $E_\lambda f$, ne diffèrent que de $\epsilon/2$ au plus. C'est-à-dire que, en désignant par $\Phi_3(\lambda)$ la fonction caractéristique de l'ensemble e_3, la valeur Df de l'opération positive $D = \Phi_{3E}(A) - \Phi_{1E}(A)$, pour l'élément f, est supposée de ne pas dépasser la quantité $\epsilon/2$. Posons encore $e_4 = e_2 - (e_2 e_3)$ et soit $\Phi_4(\lambda)$ la fonction caractéristique de l'ensemble e_4. Alors il vient par le résultat particulier que nous venons

[6] H. Lebesgue, "Sur l'intégrale de Stieltjes et sur les opérations linéaires," *Comptes Rendus Acad. Sc. Paris*, **150** (1910), pp. 86–88.

d'établir, que $\Phi_3 \parallel \Phi_4$ et par conséquent, d'après le critère formulé au commencement du chapitre III, il existe une décomposition $f = f_1 + f_2$ de sorte que

$$\Phi_3 f_1 \leqq \epsilon/2, \qquad \Phi_4 f_2 \leqq \epsilon/2.$$

Comme de plus $Df_2 \leqq Df \leqq \epsilon/2$ et $\Phi_1(\lambda) \leqq \Phi_3(\lambda)$, $\Phi_2(\lambda) \leqq \Phi_4(\lambda) + (\Phi_3(\lambda) - \Phi_1(\lambda))$, il s'ensuit que

$$\Phi_1 f_1 \leqq \epsilon/2, \qquad \Phi_2 f_2 \leqq \Phi_4 f_2 + Df_2 \leqq \epsilon$$

et d'après le critère cité cela implique que $\Phi_1 \parallel \Phi_2$.

Passons au cas de deux fonctions bornées Φ_1 et Φ_2. Or il est manifeste qu'on ne restreindra pas la généralité en supposant que les fonctions soient inférieures à l'unité et par conséquent, dominées respectivement par les fonctions caractéristiques Ψ_1 et Ψ_2 des ensembles e_1 et e_2 dans lesquels respectivement $\Phi_1 \neq 0$, $\Phi_2 \neq 0$. Il s'ensuit que l'on a aussi $\Phi_1 \leqq \Psi_1$, $\Phi_2 \leqq \Psi_2$ pour les opérations respectives. Comme de plus les ensembles e_1, e_2 sont sans point commun, les deux opérations Ψ_1 et Ψ_2 sont disjointes d'après ce que nous venons d'établir et à plus forte raison il en sera de même pour les opérations Φ_1 et Φ_2.

Enfin le cas général où les fonctions ne sont plus supposées être bornées, se ramène au cas précédant en tronquant d'abord les fonctions au niveau n et puis faisant passer n à l'infini ou ce qui revient au même, en appliquant le théorème 7 aux opérations qui correspondent aux fonctions tronquées, ce qui achève la démonstration.

Voici un *corollaire* important de notre théorème concernant les parties positive et négative d'une opération Φ correspondant à une fonction $\Phi(\lambda)$ de signe quelconque. Les fonctions $\Phi_1 = \text{pos } \Phi(\lambda)$ et $\Phi_2 = \text{neg } \Phi(\lambda)$, parties positive et négative de notre fonction, étant telles que le produit $\Phi_1\Phi_2$ s'annule partout, il vient que $\Phi_1 \parallel \Phi_2$; comme d'autre part on a $\Phi = \Phi_1 - \Phi_2$, pour les opérations de même que pour les fonctions, il s'ensuit par le théorème 16 que les opérations Φ_1 et Φ_2 ne sont que les parties positive et négative de l'opération. C'est-à-dire que *les parties positive et négative de l'opération* $\Phi_E(A)$ *sont précisément les deux opérations qui correspondent aux parties positive et négative de la fonction* $\Phi(\lambda)$.

Observons en passant que le théorème 17 reste valable pour des fonctions de signe quelconque; on n'a en effet qu'à reprendre le dernier énoncé en se rappelant ce qu'il faut entendre par opérations disjointes quand il s'agit des opérations de signe quelconque.

Le corollaire que nous venons d'énoncer suggère la généralisation suivante, embrassant à son tour le théorème 17.

THÉORÈME 18. *Étant données un nombre fini de fonctions* $\Phi_i(\lambda)$ *intégrables par rapport à* E_λ, *soient respectivement* $\underline{\Phi}(\lambda)$ *et* $\bar{\Phi}(\lambda)$ *leurs enveloppes inférieure et supérieure; alors on a*

$$\underline{\Phi}_E(A) = \inf \{\Phi_{iE}(A)\}; \qquad \bar{\Phi}_E(A) = \sup \{\Phi_{iE}(A)\}.$$

Ces faits subsistent pour une infinité dénombrable de fonctions $\Phi_i(\lambda)$, *sous la condition que l'enveloppe en question soit intégrable par rapport à* E_λ, *condition*

qui évidemment est toujours remplie pour des fonctions non-négatives $\Phi_i(\lambda)$ et pour leur enveloppe inférieure, ou aussi sous la condition que la plus grande minorante ou la plus petite majorante en question existe.

En effet, pour deux fonctions Φ_1, Φ_2, le théorème suit immédiatement de notre corollaire en rappelant que

$$\min\ (\Phi_1(\lambda), \Phi_2(\lambda)) = \Phi_1(\lambda) - \text{pos}\ (\Phi_1(\lambda) - \Phi_2(\lambda)),$$

$$\max\ (\Phi_1(\lambda), \Phi_2(\lambda)) = \Phi_2(\lambda) + \text{pos}\ (\Phi_1(\lambda) - \Phi_2(\lambda)).$$

Or le cas de n fonctions se ramène par récurrence à ce cas particulier et le cas d'une infinité dénombrable en ressort par le théorème permettant d'intervertir, au cas des suites monotones, l'ordre de l'intégration et du passage à la limite.

Le théorème 18 nous fournit les moyens pour étudier dans toute leur généralité les liens qui existent entre les fonctions numériques composées et les fonctions symboliques d'opérations qui y correspondent. Avant d'y procéder, il nous faut encore calculer les opérations qui correspondent aux fonctions $\Phi(\lambda) = 1$ et $\Phi(\lambda) = \lambda$. Je dis que ce ne sont que les opérations E et A. Quant à la première, notre assertion vient immédiatement de ce que $E_\infty = E$, $E_{-\infty} = 0$. Quant à la seconde, soit ϵ une quantité positive arbitrairement choisie et écrivons l'intégrale (1), avec $\Phi(\lambda) = \lambda$, sous la forme d'une série infinie dans les deux directions

$$\Phi = \sum_{-\infty}^{\infty} \int_{i\epsilon}^{(i+1)\epsilon} \lambda\, dE_\lambda\,;$$

écrivons de même, avec les notations du chapitre précédent,

$$A = \sum_{-\infty}^{\infty} (A_{(i+1)\epsilon} - A_{i\epsilon}).$$

Or, les termes de rang i des deux séries sont compris entre $i\epsilon(E_{(i+1)\epsilon} - E_{i\epsilon})$ et $(i+1)\epsilon(E_{(i+1)\epsilon} - E_{i\epsilon})$, ce qui est évident pour la première et s'ensuit pour la seconde des inégalités (3) du chapitre précédent. De là il vient immédiatement que

$$-\epsilon E \leqq \Phi - A \leqq \epsilon E$$

et par conséquent on a précisément $\Phi = A$.

Cela étant, nous procédons à démontrer le théorème suivant.

THÉORÈME 19. *Soit $B = \Phi_E(A)$ et soit $\bar{E}_\mu$ la décomposition de E qui correspond à l'opération B. Supposons que la fonction $\Psi(\mu)$ soit intégrable par rapport à $\bar{E}_\mu$ et posons $\Psi(\Phi(\lambda)) = \Gamma(\lambda)$. Alors cette dernière fonction est intégrable par rapport à E_λ et l'on a*

$$\Gamma_E(A) = \Psi_E(B) = \Psi_E(\Phi_E(A)).$$

Brièvement, la composition des fonctions symboliques du type $\Phi_E(A)$ répond à celle des fonctions numériques. Pour le voir, commençons par observer que si notre assertion est valable pour une fonction Ψ, il en sera de même pour la partie positive de cette fonction, ce qui vient immédiatement par le corollaire du

théorème 17. Or nous venons de voir que notre assertion est vrai pour $\Psi(\mu) = 1$ et pour $\Psi(\mu) = \mu$; donc elle est vraie pour toute fonction linéaire de μ et pour les parties positives de telles fonctions et enfin pour les combinaisons linéaires des fonctions considérées, c'est-à-dire pour toutes les fonctions continues et composées d'un nombre fini de traits linéaires. L'assertion reste aussi valable pour les limites des suites monotones, croissantes ou décroissantes, formées de fonctions du type considéré; cela vient par le théorème permettant d'intervertir l'ordre de l'intégration et du passage à la limite. De plus quand on n'envisage que des suites bornées, on est assuré d'avance que les fonctions limites sont intégrables par rapport à $\bar{E}_\mu$. Ces fonctions limites pour lesquelles on vient de vérifier le théorème, ce sont précisément toutes les fonctions bornées continues ou semicontinues; on en passe à toutes les fonctions Ψ bornées et intégrables par rapport à $\bar{E}_\mu$ en fixant, outre la fonction Ψ, un élément arbitraire f et une quantité positive ϵ et en rappelant que la fonction Ψ peut être enfermée entre deux fonctions bornées et semicontinues Ψ_1 et Ψ_2, $\Psi_1(\mu) \leqq \Psi(\mu) \leqq \Psi_2(\mu)$ et cela de sorte que la différence de l'intégrale de ces deux fonctions par rapport à $\bar{E}_\mu$ soit inférieure à ϵ. De là il vient immédiatement que

$$| \Psi_E(B)f - \Gamma_E(A)f | < \epsilon$$

et comme f et ϵ sont arbitraires, on conclut que l'assertion du théorème subsiste pour toutes les fonctions Ψ envisagées. On étend la validité de l'assertion aux fonctions positives non bornées en les tronquant d'abord au niveau n, puis faisant n aller à l'infini; enfin, on passe aux fonctions non bornées et de signe quelconque en considérant à part leurs parties positive et négative, ce qui achève la démonstration.

Toutes nos considérations restent valables, avec des modifications plus ou moins évidentes, pour les fonctions de plusieurs variables. Envisageons, pour fixer les idées, le cas de deux variables. Alors on écrira, à l'analogue de la formule (1),

$$\Phi = \Phi_E(A, B) = \int_{-\infty}^{\infty}\int \Phi(\lambda, \mu)\, dE_{\lambda,\mu} \tag{4}$$

où $E_{\lambda,\mu}$ est la décomposition à deux paramètres, envisagée à la fin du chapitre précédent, de l'opération E. La formule (4) veut indiquer que l'opération $\Phi_E(A, B)$ se calcule, pour tout élément f, en intégrant la fonction $\Phi(\lambda, \mu)$ par rapport à la fonction $E_{\lambda,\mu}f$, calcul qui se fait comme dans le cas d'une seule variable; on n'aura qu'à remplacer les différences du type $E_{\lambda_2} - E_{\lambda_1}$ par des différences itérées de la forme

$$D_{12} = E_{\lambda_1,\mu_1} + E_{\lambda_2,\mu_2} - E_{\lambda_1,\mu_2} - E_{\lambda_2,\mu_1}.$$

Or il vient de la formule (5) du chapitre précédent et des relations $E_{\lambda_2} - E_{\lambda_1} \prec E$, $\bar{E}_{\mu_2} - \bar{E}_{\mu_1} \prec E$, par le théorème 11, que ces différences itérées sont des parties disjointes de l'opération E; cela étant, le théorème 17, avec sa démonstration et

son corollaire, ainsi que le théorème 18, s'étendent immédiatement au cas de deux variables.

Le cas de plusieurs variables s'achève d'une manière analogue.

Quant au théorème 19 concernant les fonctions composées, on le généralise pour le cas de n fonctions $\Phi_1, \cdots, \Phi_n$ dépendant d'une ou de plusieurs variables $\lambda_1, \cdots, \lambda_m$ et une fonction $\Psi(\mu_1, \cdots, \mu_n)$ dans laquelle on aura à substituer aux variables μ_i les fonctions Φ_i. Pour fixer les idées, soit $m = 1$, $n = 2$, ce cas montrant nettement les modifications nécessaires dans le cas général. Nous aurons donc à considérer, outre les opérations E, A, E_λ, les opérations $B = \Phi_{1E}(A)$, $C = \Phi_{2E}(A)$ qui correspondent aux fonctions $\Phi_1(\lambda)$ et $\Phi_2(\lambda)$ et l'opération $\Psi_E(B, C)$ correspondant à la fonction $\Psi(\mu, \nu)$ et que l'on définit en intégrant la fonction Ψ par rapport à $\bar{E}_{\mu,\nu} = \inf(\bar{E}_\mu, \bar{E}_\nu)$, les décompositions $\bar{E}_\mu$ et $\bar{E}_\nu$ étant celles qui correspondent respectivement aux opérations B et C. Alors le théorème 19 reste valable, les modifications qu'il faut apporter à l'énoncé n'étant en substance que d'écrire Φ_1, Φ_2 au lieu de Φ; B, C au lieu de B et μ, ν au lieu de ν. Quant à la démonstration, la seule modification essentielle consiste en ce que, au lieu des sommes formées de fonctions linéaires et des parties positives de telles fonctions, il est préférable d'envisager des combinaisons linéaires des fonctions que l'on construit en partant d'un nombre fini de fonctions linéaires, homogènes ou non, de variables μ, ν et en passant, dans un certain ordre, à des enveloppes inférieures et supérieures. En effet, envisageons dans le plan un réseau triangulaire et formons, pour un des noeuds et pour chacun des triangles adjacents, les fonctions linéaires égales à un au noeud envisagé et s'annulant sur le côté opposé du triangle respectif. Puis, formons la partie positive de l'enveloppe inférieure de ces fonctions ou ce qui revient au même, l'enveloppe supérieure de leur enveloppe inférieure et de 0. Cela étant, envisageons une fonction continue qui est linéaire dans chacun des triangles du réseau et constante dans les triangles éloignés; alors il est manifeste qu'une telle fonction pourra être écrite toujours, à une constante additive près, sous forme d'une combinaison linéaire des fonctions particulières que nous venons de former. En observant encore que l'évaluation des intégrales

$$\int\int_{-\infty}^{\infty} d\bar{E}_{\mu,\nu}, \qquad \int\int_{-\infty}^{\infty} \mu\, d\bar{E}_{\mu,\nu}, \qquad \int\int_{-\infty}^{\infty} \nu\, d\bar{E}_{\mu,\nu}$$

par intégration successive, exécutée pour les deux dernières respectivement dans l'ordre ν, μ et dans l'ordre μ, ν, donne $\Psi_E(B, C) = E, B, C$ respectivement pour $\Psi(\mu, \nu) = 1, \mu, \nu$, il vient que le théorème 19 est valable pour toutes les fonctions que nous venons d'envisager. En faisant varier notre réseau, par exemple en le subdivisant indéfiniment d'une manière convenable, nous disposerons d'une classe de fonctions dont les suites monotones suffisent largement pour approcher indéfiniment toutes les fonctions continues ou semicontinues et bornées des variables μ, ν; cela étant, la démonstration s'achève comme pour les fonctions d'une seule variable.

A titre d'exemple, disons quelques mots du *produit* des opérations A et B, formé par rapport à l'opération E et que nous désignerons par $(AB)_E$; les opérations A et B sont supposées d'appartenir à la famille de E ou à son prolongement. Nous entendons par là l'opération $\Phi_E(A, B)$ qui correspond au produit $\Phi(\lambda, \mu) = \lambda\mu$. Comme l'intégrale de cette fonction non bornée par rapport aux fonctions $E_{\lambda,\mu}f$ ne converge pas nécessairement dans tous les cas, le symbole $(AB)_E$ n'a pas toujours de sens; mais on voit immédiatement que le produit existe, entre autres, lorsque l'une des opérations A, B est bornée par rapport à E, c'est-à-dire qu'elle est comprise entre des multiples numériques de E.

La multiplication que nous venons d'introduire est évidemment commutative; les autres règles de calcul, moins évidentes, sont des corollaires du théorème 19. Tels sont la loi distributive $(A(B + C))_E = (AB)_E + (AC)_E$, valable sous l'hypothèse que deux des trois produits impliqués aient un sens et la loi associative $((AB)_E C)_E = (A(BC)_E)_E$, soumise à l'hypothèse que les produits $(AB)_E$, $(BC)_E$ et l'un au moins des deux produits itérés existent.

Dans mon mémoire original où j'ai étudié ces produits d'opérations d'une façon détaillée, je les ai introduits sans faire appel à l'intégrale (4). Dans le cas où A et B sont des opérations positives, cas auquel on remène sans peine le cas général, ma définition d'autrefois était la suivante. Le produit $(AB)_E$ se définit comme la plus petite majorante de toutes les opérations du type abE' où E' parcourt les parties disjointes de l'opération E et les facteurs numériques $a \geqq 0$, $b \geqq 0$ varient sous les conditions $aE' \leqq A$ et $bE' \leqq B$. Je laisse au lecteur de vérifier l'équivalence des deux définitions.

Voici un autre exemple, plus intéressant par des raisons que nous allons voir. Supposons que $A \geqq 0$, $B \geqq 0$ et considérons l'opération $(A^{\frac{1}{2}}B^{\frac{1}{2}})_E$ qui correspond à la fonction $\lambda^{\frac{1}{2}}\mu^{\frac{1}{2}}$, envisagée pour $\lambda \geqq 0$, $\mu \geqq 0$; il est clair que sous l'hypothèse faite, les valeurs négatives des variables ne comptent pas, c'est-à-dire que dans la définition (4) de $\Phi_E(\lambda, \mu)$, il est permis de remplacer par zéro la limite inférieure d'intégration, nos théorèmes restant applicables après cette modification. Cela étant, je dis que *la "moyenne géométrique"* $(A^{\frac{1}{2}}B^{\frac{1}{2}})_E$ ou brièvement $A^{\frac{1}{2}}B^{\frac{1}{2}}$ *ne dépend pas du choix particulier de l'opération* E. En effet, la fonction $\lambda^{\frac{1}{2}}\mu^{\frac{1}{2}}$ n'est que l'enveloppe inférieure des fonctions $(c\lambda + \mu/c)/2$ où c parcourt une suite partout dense de valeurs positives, par exemple celle des nombres positifs rationnels. Il s'ensuit par le théorème 18 que l'opération $A^{\frac{1}{2}}B^{\frac{1}{2}}$ n'est que la plus grande minorante des opérations $(cA + B/c)/2$ où c parcourt les nombres positifs rationnels ou, ce qui revient au même, toutes les valeurs positives, c'est-à-dire que notre "moyenne géométrique" peut être construite sans faire appel à l'opération E. D'ailleurs, en rappelant la démonstration du théorème 1, on aboutit aisément à la construction suivante. La valeur de $A^{\frac{1}{2}}B^{\frac{1}{2}}$ pour un élément donné f est la borne inférieure de l'ensemble des valeurs que prend l'expression

$$\sum_1^n (Af_k Bf_k)^{\frac{1}{2}}$$

pour tous les n et pour toutes les décompositions $f = f_1 + \cdots + f_n$ de l'élément f.

Ce dernier exemple et d'autres plus généraux, comme $A^\alpha B^\beta$ avec $\alpha, \beta > 0$, $\alpha + \beta = 1$ ou $A^\alpha B^\beta C^\gamma$ avec $\alpha, \beta, \gamma > 0$, $\alpha + \beta + \gamma = 1$, nous suggèrent d'envisager l'hypothèse générale sous laquelle l'opération Φ_E correspondant à la fonction Φ est indépendante du choix particulier de l'opération E. Pour fixer les idées, nous envisageons les fonctions de deux variables et cela en observant que le cas des fonctions d'une seule variable est dépourvu d'intérêt, n'embrassant que le cas des fonctions de la forme $a\lambda + b\,|\,\lambda\,|$. Nous allons montrer que *l'opération Φ_E est indépendante de E toujours que la fonction $\Phi(\lambda, \mu)$ est positivement homogène du premier ordre*, c'est-à-dire que, pour tout $c \geqq 0$,

$$\Phi(c\lambda, c\mu) = c\Phi(\lambda, \mu). \tag{5}$$

A cet effet, observons qu'une large classe de telles fonctions s'obtient en partant des fonctions linéaires homogènes ou ce qui revient au même, des variables λ, μ elles-mêmes et en formant alternativement, de toutes les manières possibles, des enveloppes et des combinaisons linéaires. On arrive aux même fonctions en partageant le plan en un nombre fini de régions angulaires issues de l'origine et en considérant les fonctions continues dans le plan entier et égales, dans chacune de ces régions, à une fonction linéaire homogène dont l'expression varie de région à région. Or on voit immédiatement que ces fonctions suffisent pour approcher indéfiniment, par des suites monotones, toutes les fonctions semi-continues Φ du type (5) et en particulier celles qui sont inférieures en module à un multiple numérique de la fonction $|\,\lambda\,| + |\,\mu\,|$. Donc pour toutes ces fonctions notre assertion se vérifie par le théorème 18, énoncé pour des fonctions de deux variables. C'est-à-dire que pour toutes ces fonctions l'opération $\Phi_E(A, B)$ existe, au sens de la formule (4), quelle que soil l'opération linéaire positive E, sous la seule hypothèse que A et B appartiennent à la famille de E ou à son prolongement et de plus, la fonction $\Phi(\lambda, \mu)$ étant fixée, l'opération $\Phi_E(A, B)$ est la même pour toutes les E. Soient maintenant $\bar{E}$ et $\bar{\bar{E}}$ deux quelconques parmi les opérations E envisagées, soient $\bar{E}_{\lambda,\mu}$ et $\bar{\bar{E}}_{\lambda,\mu}$ leurs décompositions par rapport à A et B et supposons que la fonction $\Phi(\lambda, \mu)$, du type (5), soit intégrable par rapport à la décomposition $\bar{E}_{\lambda,\mu}$; en attendant, nous supposons encore que $|\,\Phi(\lambda, \mu)\,|$ reste inférieur à un multiple de $|\,\lambda\,| + |\,\mu\,|$. Soient de plus ϵ une quantité positive et f un élément arbitrairement choisis; alors en raisonnant comme au cours de la démonstration du théorème 19, nous enfermons Φ entre deux de nos fonctions semicontinues soit Φ_1 , Φ_2 où $\Phi_1(\lambda, \mu) \leqq \Phi(\lambda, \mu) \leqq \Phi_2(\lambda, \mu)$, choisies de sorte que les intégrales de ces deux fonctions par rapport à $\bar{E}_{\lambda,\mu} f$ ne différent que de ϵ au plus. D'après ce que nous venons de voir, la valeur de ces intégrales n'est pas changée quand on remplace $\bar{E}$ par $\bar{\bar{E}}$; comme de plus f et ϵ étaient arbitraires, il vient que $\Phi_{\bar{E}}(A, B) = \Phi_{\bar{\bar{E}}}(A, B)$.

Enfin, pour les autres fonctions Φ positivement homogènes et intégrables par rapport à la décomposition $E_{\lambda,\mu}$, la démonstration s'achève en considérant à part les parties positives et négatives de ces fonctions et en les tronquant au niveau n ($|\,\lambda\,| + |\,\mu\,|$), c'est-à-dire en formant, par exemple, l'enveloppe inférieure de pos Φ et de $n(|\,\lambda\,| + |\,\mu\,|)$, et en faisant passer n à l'infini.

Pour illustrer ces considérations, envisageons l'exemple 1 du chapitre I,

c'est-à-dire le domaine formé par les fonctions continues $f(x)$ définies dans l'intervalle $a \leqq x \leqq b$. On sait que les opérations linéaires de ce domaine Ω peuvent être écrites sous la forme d'une intégrale de Stieltjes

$$Af = \int_a^b f(x)\, d\alpha(x), \tag{6}$$

où $\alpha(x)$, la "génératrice" de l'opération A, est une fonction à variation bornée que l'on peut supposer continue à droite à l'intérieur de l'intervalle (a, b) et s'annulant pour $x = a$, hypothèses sous lesquelles la fonction $\alpha(x)$ sera uniquement déterminée par l'opération A.[7] Inversement, chaque fonction à variation bornée donne lieu à une opération A. Les opérations positives sont précisément celles dont les génératrices sont des fonctions non-décroissantes. Soit E une telle opération et soit $\epsilon(x)$ sa génératrice; quant à l'opération A, supposons d'abord qu'elle soit positive et bornée par rapport à E, $0 \leqq A \leqq cE$; alors, grâce à un théorème de M. Lebesgue, on pourra écrire, au lieu de (6),

$$Af = \int_a^b a(x) f(x)\, d\epsilon(x) \tag{7}$$

où la fonction $a(x)$, positive et bornée, n'est que la dérivée de $\alpha(x)$ par rapport à $\epsilon(x)$, existant presque partout par rapport à cette dernière et où l'intégrale au second membre est prise au sens de Stieltjes-Lebesgue. Dans le cas général où l'opération A, de signe quelconque et bornée ou non par rapport à E, est supposée d'appartenir à la famille de E ou à son prolongement, on n'aura qu'à considérer l'opération $A_\lambda - A_0$, positive et bornée par rapport à E conformément à l'inégalité (3) du chapitre précédent, et allant en croissant respectivement vers pos A et neg A pour $\lambda \to \pm \infty$. Donc on n'aura qu'à écrire l'opération $A_\lambda - A_0$ sous la forme (7), avec une fonction $a(x; \lambda)$ au lieu de $a(x)$ et faire les deux passages à la limite; alors la formule (7) sera vérifiée dans le cas général, avec

$$a(x) = \lim_{\lambda \to \infty} [a(x; \lambda) - a(x; -\lambda)].$$

Enfin, le fait que cette fonction $a(x)$, intégrable par rapport à $\epsilon(x)$, n'est que la dérivée de $\alpha(x)$ par rapport à $\epsilon(x)$, vient en comparant les formules (6) et (7).

Observons encore les faits évidents que l'opération $c_1A + c_2E$ appartient aussi au type envisagé et qu'elle s'écrit en remplaçant, au second membre de (7), $a(x)$ par $c_1a(x) + c_2$ et que de plus si B est une opération du même type (pour laquelle le multiplicateur $a(x)$ est à remplacer par $b(x)$), il en est de même pour les combinaisons linéaires de A, B et E et les multiplicateurs respectifs sont fournis par les combinaisons analogues des fonctions $a(x)$, $b(x)$ et 1. De plus les opérations inf (A, B) et sup (A, B) appartenant à la famille de E ou à son prolongement, sont nécessairement du type envisagé et il vient immédiatement que les multiplicateurs respectifs ne sont que les enveloppes, inférieure et supé-

[7] F. Riesz, "Sur les opérations fonctionnelles linéaires," *Comptes Rendus Acad. Sc. Paris*, **149** (1909), pp. 947-977.

rieure, des fonctions $a(x)$, $b(x)$. Il ressort de ces remarques, d'abord pour les fonctions $\Phi(\lambda)$ et $\Phi(\lambda, \mu)$ continues et composées d'un nombre fini de traits linéaires, que les multiplicateurs qui correspondent aux opérations $\Phi_E(A)$ et $\Phi_E(A, B)$ ne sont que les fonctions composées $\Phi(a(x))$ et $\Phi(a(x), b(x))$, et on démontre aisément, en raisonnant comme nous venons de le faire à d'autres occasions, que le même fait subsiste pour toutes les fonctions Φ intégrables respectivement par rapport à E_λ ou $E_{\lambda,\mu}$.

Un cas particulier très simple se présente lorsque E est l'intégration

$$Ef = \int_a^b f(x)\, dx,$$

donc $\epsilon(x) = x$, $a(x) = \alpha'(x)$ et toutes les intégrales qui interviennent sont des intégrales de Lebesgue. Un autre cas particulier est celui où $\epsilon(x)$ est une fonction des sauts; dans ce cas les intégrales se réduisent à des sommes finies ou à des séries absolument convergentes.

IX. L'extension du domaine fondamental

Un des cas particuliers les plus importants d'une telle extension se présente quand, pour introduire l'intégrale de Lebesgue, on passe des fonctions continues ou encore de celles du type escalier aux fonctions sommables, passage que l'on peut effectuer de plusieurs manières, entre autres en se servant de suites bornées ou aussi de suites monotones de fonctions, convergeant partout ou seulement presque partout; on peut aussi envisager des suites qui, au lieu de converger au sens ordinaire, ne sont supposées que de converger en mesure ou en moyenne ou seulement d'une façon faible par rapport à tel et tel exposant. Pour procéder d'une manière analogue quand il s'agit des domaines abstraits, il faudrait commencer par fixer une notion de convergence pour les suites formées d'éléments du domaine envisagé. Une des définitions possibles, d'ordre général, est la suivante. La suite $\{f_n\}$ est dite convergente si les suites numériques $\{Af_n\}$ convergent quelle que soit l'opération linéaire A. Des définitions moins restrictives se posent en exigeant la convergence des suites $\{Af_n\}$ seulement pour une certaine famille d'opérations. Ainsi par exemple, quand il s'agit du domaine Ω formé par les fonctions continues et non négatives dans l'intervalle $a \leqq x \leqq b$, la première hypothèse revient à considérer comme convergentes celles des suites qui sont bornées et convergent partout, tandis que les définitions moins restrictives dont nous venons de parler, embrassent les suites bornées ou non qui convergent partout ainsi que certains types de convergence faible.

Un principe d'extension plus systématique nous est suggéré par l'idée de dualité. Nous nous contentons d'esquisser quelques idées vagues qui s'y rattachent. Nous partons du fait que les opérations linéaires et positives d'un domaine Ω forment à leur tour un domaine Ω^* satisfaisant aux mêmes hypothèses, savoir aux hypothèses 1–4 posées au chapitre 1 et même à l'hypothèse 4′. Parmi les opérations linéaires A^* de ce nouveau domaine fondamental se trouvent, dans un sens à préciser, les éléments f du domaine Ω dont nous sommes

partis. En effet, un tel élément f fait correspondre à chaque opération A, c'est-à-dire à chaque élément A du domaine Ω^*, la valeur Af et l'opération ainsi définie dans le domaine Ω^* est évidemment linéaire. Cela suggère de faire l'extension du domaine Ω en y ajoutant toutes les autres opérations linéaires du domaine Ω^*. Or considérons sous cet aspect notre exemple 1, le domaine Ω formé par les fonctions $f(x)$, positives et continues pour $a \leqq x \leqq b$. Nous savons déjà que, dans le cas envisagé, les opérations A sont de la forme

$$Af = \int_a^b f(x)\, d\alpha(x)$$

où $\alpha(x)$ est une fonction non décroissante, d'ailleurs arbitraire et que de cette sorte, le domaine Ω^* pourra être considéré, en substance, comme formé par ces fonctions $\alpha(x)$, toujours en convenant de regarder comme identiques deux fonctions $\alpha(x)$ lorsqu'elles donnent lieu à la même opération A, ce qui revient à ce que les deux fonctions ne diffèrent que d'une constante, sauf peut-être en un ensemble dénombrable de points x, intérieurs à l'intervalle (a, b). Or parmi les opérations linéaires que l'on peut définir dans ce domanie Ω^*, se trouvent entre autres celles de la forme

$$A^*\alpha = \sum_{a<x<b} g(x)[\alpha(x+0) - \alpha(x-0)]$$

où $g(x)$ désigne une fonction bornée, d'ailleurs quelconque, mesurable ou non. Ces opérations particulières n'agissent que sur la partie de $\alpha(x)$ connue sous le nom fonction des sauts; elles peuvent aussi être écrites sous forme d'intégrale de Stieltjes-Lebesgue pour les fonctions $\alpha(x)$ qui coïncident avec leurs fonctions des sauts, mais elles ne le peuvent pas pour les fonctions $\alpha(x)$ du type général. Tout cela montre que l'extension envisagée est trop large sous certain aspect, de sorte qu'il est préférable de se limiter, au lieu de toutes les opérations linéaires du domaine Ω^*, à des opérations A^* qui, pour ainsi dire, sont assez proches de celles attachées aux éléments f du domaine Ω. A cet effet, considérons la plus petite classe d'opérations A^* embrassant celles attachées aux éléments f et jouissant des deux propriétés suivantes: 1) elle forme un module, c'est-à-dire que, avec A_1^* et A_2^*, elle contient aussi la différence $A_1^* - A_2^*$; 2) elle est complète, c'est qu'elle comprend les plus petites majorantes de ses sous-ensembles majorables. L'existence d'une telle classe est immédiate; en effet, elle n'est que la partie commune de toutes les classes jouissant des propriétés exigées. Or envisageons une opération A^* appartenant à la classe considérée et attachons à cette opération la fonction $a^*(x)$ définie pour chaque x_0 par les formules

$$a^*(x_0) = A^*(A_{x_0}); \qquad A_{x_0} f = f(x_0);$$

on voit aisément que ces fonctions $a^*(x)$ donnent précisément l'ensemble des fonctions bornées et mesurables (B) dans l'intervalle (a, b) et que de plus l'opération A^* s'exprime par l'intégrale de Stieltjes-Lebesgue

$$A^*\alpha = \int_a^b a^*(x)\, d\alpha(x). \tag{1}$$

Appliquons ces considérations à l'exemple 7 dont l'exemple 1 n'est qu'un cas particulier et qui nous conduira à la théorie d'intégration due à M. Daniell. Sans restreindre la généralité, nous pouvons supposer que, tout comme dans l'exemple 1, le domaine Ω comprenne les multiples $cf(x)$, $c \geqq 0$, de ses éléments $f(x)$. En effet, s'il n'en était pas ainsi, on pourrait tout d'abord ajouter au domaine Ω, d'une façon évidente, les multiples $cf(x)$ où c est rationnel et il est aussi manifeste que l'hypothèse 4′ subsiste après cette extension. De plus comme on aura défini l'opération A pour les nouveaux éléments cf en posant $A(cf) = cAf$, le domaine Ω^* ne sera pas changé. Une seconde extension, faisant entrer aussi les c irrationnels (et quelquefois, comme le montrent les exemples 2 et 3, d'autres éléments nouveaux) se fera en ajoutant encore les limites des suites qui convergent d'une manière relativement uniforme, comme les a définies E. H. Moore, c'est-à-dire des suites $f_m(x)$ telles que, pour presque tous les m et n

$$|f_m(x) - f_n(x)| \leqq \epsilon f_0(x)$$

où ϵ est une quantité positive arbitrairement fixée et $f_0(x)$ un élément de Ω attaché à la suite.[8] On voit aisément que le domaine ainsi étendu remplit les hypothèses 1–4 et aussi 4′ et que les opérations A du domaine Ω se prolongent au domaine étendu.

Supposons donc que le domaine Ω comprenne avec $f(x)$ tous les multiples $cf(x)$, $c \geqq 0$; alors nous voici au point de départ de la théorie de M. Daniell. Cela étant, il ne faut que suivre le même ordre d'idées que dans le cas des fonctions continues pour arriver au domaine Ω^*, ensemble des opérations A de Ω, et au plus petit module complet d'opérations A^* embrassant les opérations particulières qui correspondent aux éléments $f(x)$ de Ω et aux fonctions $a^*(x)$ introduites comme plus haut. Mais il convient de noter que dans le cas de l'exemple 1 la fonction $a^*(x)$ attachée à l'opération A^*, une fois construite, détermine l'opération A^* conformément à la formule (1). Or il n'est pas ainsi dans le cas général; pour le voir, nous n'avons qu'à modifier l'exemple 1 en limitant la variable aux valeurs rationnelles de x, mais, en récompense, en exigeant des fonctions $f(x)$ d'être uniformément continues. Ces conventions un peu artificielles semblent de ne rien changer et on ne se rend compte d'une différence essentielle entre les deux exemples que lorsqu'on arrive aux fonctions $a^*(x)$. Il est manifeste que ces fonctions ne déterminent plus les opérations A^* pour tous les éléments A du domaine Ω^*, ni même pour les A du type particulier $A_{x_0}f = f(x_0)$ sauf quand x_0 est rationnel.

Ce phénomène explique pourquoi la théorie de M. Daniell était forcée à ne considérer comme intégrale que celles des opérations A qui satisfont encore à l'hypothèse (L), savoir que $Af_n \to 0$ pour toute suite $f_n(x)$ décroissante et tendant partout vers zéro. Dans l'exemple 1, l'hypothèse (L) est réalisée pour toute opération A, comme cela vient par le théorème classique de Dini; on voit immédiatement qu'il n'en est pas ainsi pour l'exemple modifié.

---

[8] E. H. Moore, "On the Foundations of the Theory of Linear Integral Equations," *Bulletin American Math. Soc.*, **13** (1912), pp. 334–362.

Montrons que *les opérations A jouissant de la propriété* (L) *forment une famille complète.* Il est évident que ces opérations satisfont aux deux premières hypothèses posées pour les familles complètes; avec A_1, elles comprennent aussi toutes les $A \leqq A_1$ et avec A_1 et A_2, leur somme $A_1 + A_2$. Ce qu'il faut encore prouver, c'est que la plus petite majorante de tout sous-ensemble majorable appartient aussi à la classe considérée. C'est-à-dire que si $Af_n \to 0$ pour toute suite décroissante $f_n(x) \to 0$ et pour tout élément d'un ensemble majorable $\{A\}$, on a aussi $\bar{A}f_n \to 0$ pour $\bar{A} = \sup \{A\}$.

Fixons une quantité $\epsilon > 0$; alors, d'après la construction de $\sup \{A\}$, on peut tirer de notre ensemble un nombre fini d'opérations, soit $A_1, A_2, \cdots, A_m$ et faire une décomposition $f_1 = f_{11} + f_{12} + \cdots + f_{1m}$ de façon que

$$\sum_1^m A_k f_{1k} \geqq \bar{A}f_1 - \epsilon$$

et par conséquent, en posant $\sup (A_1, A_2, \cdots, A_m) = \bar{A}_m$, on aura $\bar{A}_m f_1 \geqq \bar{A}f_1 - \epsilon$, donc $(\bar{A} - \bar{A}_m)f_1 \leqq \epsilon$; comme de plus $\bar{A} - \bar{A}_m \geqq 0$, $f_n \leqq f_1$, il vient que, pour tout n,

$$(\bar{A} - \bar{A}_m)f_n \leqq \epsilon. \tag{2}$$

D'autre part on a

$$\bar{A}_m f_n \leqq \sum_{k=1}^m A_k f_n$$

donc par (2), pour $n \to \infty$,

$$\overline{\lim}\ \bar{A}f_n = \overline{\lim}\ (\bar{A} - \bar{A}_m)f_n \leqq \epsilon,$$

c'est que $\bar{A}f_n \to 0$, C. Q. F. D.

Pour élucider complètement l'hypothèse (L) au point de vue de notre théorie, il nous faudrait encore examiner si en nous bornant, au lieu de toutes les opérations A, à celles qui satisfont à l'hypothèse (L) et en nous servant des opérations A^* de ce domaine réduit considéré comme domaine fondamental, l'extension ainsi faite du domaine Ω n'est pas plus large que celle basée sur le domaine Ω^* et si, d'autre part, après cette réduction, les opérations A^* sont déterminées univoquement par les fonctions $a^*(x)$ qui leur correspondent. La réponse à la dernière question sera en général négative, comme le montre déjà l'exemple 1, et pour arriver à une extension intrinsèque qui n'est pas excessivement large, le mieux sera, dans la plupart des cas, d'avoir recours au plus petit module complet embrassant les opérations A^* qui correspondent aux fonctions $f(x)$ dont nous sommes partis.

SZEGED, HUNGARY.

ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

CHARACTER GROUP OF A RELATIVE HOMOLOGY GROUP

BY WILLIAM W. FLEXNER

(Received April 21, 1939)

1. Introduction

If L is a subcomplex of a finite complex K, the topologically invariant factor group of the group of p-cycles of L bounding in K taken with respect to the subgroup of those bounding in L is useful in topology. The principal aim of this paper is to find its character group and to interpret the character group geometrically (theorems 4.2 and 4.4 at the end). Incidentally some algebraic and topological side results are obtained which will, it is hoped, help in finding character groups of less simple topologically invariant groups.

The set of elements of a group containing more than one element are turned into a Hausdorff space by definition of open sets which satisfy, besides the Hausdorff axioms, axioms connecting the group operation and open sets, for instance those given by van Kampen[1] or Pontrjagin.[2] A *neighborhood* of an element of the group is an open set containing the element, and a *nucleus* is a neighborhood of the identity. The *closure* of a set F of a Hausdorff space will be written $\overline{F}$. A Hausdorff space is *bicompact* if every covering of the whole space by open sets has a finite subset which covers the whole space (A.H. p. 86[3]). The space will be called *discrete* if every element is an open set.

As all groups here considered are Abelian, the group operation will be written additively, $a + b$. If H is the group of continuous homomorphisms of the topological Abelian group G into (A.H. "in") the additive group R of the real numbers mod 1 with the usual topology, then H is called the *character group* of G. When G is discrete and H is topologized as in v.K. or P.B., H is bicompact. When G is bicompact, H is given discrete topology.

If G and H are topological Abelian groups and to every $g \in G$ and $h \in H$ there is an element $gh = hg$ of R such that gh is a continuous function of g and of h, $(g + g')\,h = gh + g'h$ and $g(h + h') = gh + gh'$, then G and H are said to be *paired*; each h defines a continuous homomorphism of G into R, and each g a continuous homomorphism of H into R. If G and H are paired and $\mathfrak{H}$ is a subgroup of H, $\mathfrak{G} = (G, \mathfrak{H})$ is defined as the set of all elements $g \in G$ such that $gh = 0$ for every $h \in \mathfrak{H}$. The set $(G, \mathfrak{H})$ is a subgroup of G. (P.B. Note that none of the results in this paper depend on the Weyl-Peter theorem, see P.G.[4])

[1] van Kampen: *Locally Bicompact Abelian Groups and their Character Groups;* Annals of Math. 36 (1935) pp. 448–463. Referred to in the sequel as "v.K.".

[2] Pontrjagin: *Topological Groups,* (translation by E. Lehmer) Princeton, 1939. Referred to as "P.B."

[3] Alexandroff-Hopf: *Topologie,* Berlin 1935. Referred to as "A.H."

[4] Pontrjagin: *The Theory of Topological Commutative Groups:* Annals of Math. (35) 1934, pp. 361–388, referred to in the sequel as "P.G."

A. The necessary and sufficient condition that a discrete group G and a bicompact group H be character groups of one another is $(G, H) = (H, G) = 0$. (P.B.; P.G. p. 369; v.K. pp. 445, 456)

If $\mathfrak{G}$ is a closed subgroup of G (i.e. a subgroup whose elements form a closed set) the factor group $G - \mathfrak{G}$ may be topologized by calling open in $G - \mathfrak{G}$ any set of $G - \mathfrak{G}$ corresponding to an open set in G. (v.K. p. 450)

B. Then if G is discrete (bicompact) and $\mathfrak{G}$ is a closed subgroup of G, $G - \mathfrak{G}$ is discrete (bicompact). (v.K. p. 450)

By sG is meant the direct sum to s terms (s finite)

$$sG = G + G + \cdots + G.$$

If g^i is an arbitrary element of G, the typical element of sG will be written in vector form as $g = (g^1, \cdots, g^s)$. To topologize sG call open in sG any set U of elements g such that $g^i \subset U^i$, $i = 1, \cdots, s$, where U^i is open in G: $U = (U^1, \cdots, U^s)$. To get the group operation in sG, if $g' = (g'^1, \cdots, g'^s)$ then $g + g' = (g^1 + g'^1, \cdots, g^s + g'^s)$.

C. If G is discrete, sG is discrete. If G is bicompact, sG is bicompact.

D. If G and H are character groups and therefore paired, sG and sH can be paired by the rule

$$gh = \sum_{i=1}^{s} g^i h_i$$

where $h = (h_1, \cdots, h_s)$, $h_i \in H$, is the typical element of sH. It follows from A and B that sG and sH are character groups.

E. If G is discrete or bicompact, H is its group of characters and $\mathfrak{H}$ is a closed subgroup of H, $\mathfrak{G} = (G, \mathfrak{H})$ and $\mathfrak{H} = (H, \mathfrak{G})$ are equivalent properties (P.B.; P.G.).

F. If F is a transformation from sG to tG such that

$$F(g^1, \cdots, g^s) = (f^1, \cdots, f^t)$$

where $f^i = \sum_{j=1}^{s} a_{ij} g^j$, a_{ij} integers, then F is continuous.

G. The subgroup $A = \{g\}$ of sG such that $Fg = 0$ is closed in sG. This follows from F and the fact that 0 is closed in tG.

H. The set of elements Fg, $g \in sG$, is closed in tG, G bicompact. This is because every continuous transformation of a bicompact space into a Hausdorff space carries closed sets into closed sets (A.H. p. 95) and sG is closed.

2. Algebraic lemmas

The following notation is now useful:

a. $\mathfrak{G} < G$ *means* $\mathfrak{G}$ *is a subgroup of* G.

b. $\mathfrak{G} \ll G$ *means* $\mathfrak{G}$ *is a closed subgroup of* G.

c. *When* $F < sH$, F^σ *shall be the set of elements* $(f_1, \cdots, f_\sigma)$ *of* σH *for which* $(f_1, \cdots, f_\sigma, 0, \cdots, 0)$ *is an element of* F.

d. *When* $F < sH$, F^* *is the set of elements* $(f_1, \cdots, f_\sigma)$ *of* σH *such that* $(f_1, \cdots, f_\sigma, \cdots, f_s) \in F$.

e. *When* $F < \sigma H$, $[F]$ *is the set of elements* $(f_1, \cdots, f_\sigma, \cdots, f_s)$ *such that* $(f_1, \cdots, f_\sigma) \in F$ *and* $f_{\sigma+i}$ *is arbitrary in* H.

f. *When* $F < sH$, $[F] = [F^*]$

g. $\mathfrak{G} \approx \mathfrak{H}$ *means* $\mathfrak{G}$ *and* $\mathfrak{H}$ *are isomorphic.*

h. *If* $A < G$, $B < G$ *then* $A \smile B$ *is the set of all elements* $a + b$ *in* G, $a \in A$, $b \in B$. *(Obviously* $A \smile B < G$.)

It is apparent that $F < sH$ implies $F^\sigma < F^* < \sigma H$ and that $F < \sigma H$ or $F < sH$ implies $[F] < sH$. Immediate lemmas are:

1. *If* $F \ll sG$ *then* $F^\sigma \ll \sigma G$.
2. *If* G *is discrete or bicompact and* $A, B, \ll G$, *then* $A \smile B \ll G$.
3. *If* H *is discrete or bicompact and* $F \ll sH$, *then* $F^\sigma \ll F^* \ll \sigma H$.
4. *If* $F \ll \sigma H$ *or* $F \ll sH$ *and* H *is discrete or bicompact, then* $[F] \ll sH$.

THEOREM 2.1. *If* G *and* H *are character groups, one being discrete, the other bicompact and if* $F < sH$, *then* $(sG, F)^\sigma = (\sigma G, F^*)$.

Let $g = (g^1, \cdots, g^\sigma, \cdots, g^s)$ and $g' = (g^1, \cdots, g^\sigma)$; $f = (f_1, \cdots, f_s)$. If $g' \in (\sigma G, F^*)$ then $\sum_{i=1}^{\sigma} g^i f_i = 0$ for every $f \in F$, and $\sum_{i=1}^{s} g^i f_i = \sum_{i=\sigma+1}^{s} g^i f_i$ for every $f \in F$. So for $g'' = (g^1, \cdots, g^\sigma, 0, \cdots, 0)$ $\sum_{i=1}^{s} g^i f_i = 0$ for every $f \in F$. Hence g'' is in (sG, F) and the corresponding g' therefore is in $(sG, F)^\sigma$, so $(sG, F)^\sigma \supset (\sigma G, F^*)$.

If $g' \in (sG, F)^\sigma$ then $\sum_{i=1}^{\sigma} g^i f_i = 0$ for every $f \in F$ and hence for every $f \in F^*$. Therefore $g' \in (\sigma G, F^*)$ and $(\sigma G, F^*) \supset (sG, F)^\sigma$ which completes the proof.

THEOREM 2.2. *If* G *and* H *are character groups and* $F < \sigma G$, *then,* $(sG, [F]) \approx (\sigma G, F)$.

Let I be the isomorphism $I(g^1, \cdots, g^\sigma, 0, \cdots, 0) = (g^1, \cdots, g^\sigma)$. If $(g^1, \cdots, g^\sigma) \in (\sigma G, F)$ then $\sum_{i=1}^{\sigma} g^i f_i = 0$ for every $(f_1, \cdots, f_\sigma) \in F$, and $I^{-1}(g^1, \cdots, g^\sigma) \in (sG, [F])$. On the other hand, if $(g^1, \cdots, g^s) \in (sG, [F])$

$$\sum_{i=1}^{\sigma} g^i f_i + \sum_{j=\sigma+1}^{s} g^j f_j = 0, \tag{1}$$

where $(f_1, \cdots, f_\sigma) \in F$ and $f_{\sigma+1}, \cdots, f_s$ are arbitrary elements of H. Setting $f_{\sigma+1} = \cdots = f_s = 0$ gives

$$\sum_{i=1}^{\sigma} g^i f_i = 0, \tag{2}$$

so (1) implies $\sum_{j=\sigma+1}^{s} g^j f_j = 0$ for every choice of $f_{\sigma+1}, \cdots, f_s$. Suppose that $g^{\sigma+i} \neq 0$ for the element $g = (g^1, \cdots, g^s)$ of $(sG, [F])$. Then by A there is an $f'_{\sigma+i} \in H$ such that $g^{\sigma+i} f'_{\sigma+i} \neq 0$. As element of $[F]$ choose $f = (f_1, \cdots, f_s)$ where $(f_1, \cdots, f_\sigma) \in F$ and $f_j = \delta_j^{\sigma+i} f'_{\sigma+i}$, $\sigma < j \leqq s$. Then by (2), $gf = g^{\sigma+i} f'_{\sigma+i} \neq 0$ and g is not in $(sG, [F])$. So all elements of $(sG, [F])$ are of the form $(g^1, \cdots, g^\sigma, 0, \cdots, 0)$ with $\sum_{i=1}^{\sigma} g^i f_i = 0$ by (2). Hence $I^{-1}(\sigma G, F) \supset (sG, [F])$.

THEOREM 2.3. *If* $F < sH$ *then* $(sG, F)^\sigma \approx (sG, [F])$.

By theorems 2.1 and 2.2 and the definition of $[F]$, $(sG, F)^{\sigma} \approx (\sigma G, F^*) \approx (sG, [F^*]) = (sG, [F])$.

THEOREM 2.4. *If $A, B < G$, then $G - A \smile B = (G - A) - B^A$ where B^A is the set of elements of $G - A$ (cosets) such that each coset in B^A contains an element of B.*

The proof depends on showing that the necessary and sufficient condition that two elements, g and g' of G be in the same coset of $(G - A) - B^A$ is that $g - g' \in A \smile B$.

THEOREM 2.5. *If G and H are character groups and $A, B < G$, $C, D < H$ and 1. $A = (G, C)$, 2. $B = (G, D)$; then $A \frown B = (G, C \smile D)$, where $A \frown B$ is the set common to A and B.*

Suppose $x \in A \frown B$ and $c \in C$, $d \in D$. Then $x(c + d) = xc + xd = 0$ by (1) and (2) since $x \in A$ and $x \in B$. So $x \in (G, C \smile D)$ and $A \frown B \subset (G, C \smile D)$. Suppose $x \in (G, C \smile D)$. Then $x(c + d) = 0$ for every $c \in C$ and $d \in D$. Take $d = 0$; then $xc = 0$ for every $c \in C$ and so by (1) $x \in A$. Similarly $x \in B$ so $x \in A \frown B$ and $A \frown B \supset (G, C \smile D)$.

THEOREM 2.6. *If G and H are character groups of one another, one being discrete, the other bicompact and*

$$\text{i.} \quad \Gamma \ll \mathfrak{G} \ll G$$

$$\theta \ll \mathfrak{H} \ll H$$

$$\text{ii.} \quad \mathfrak{G} = (G, \theta) \tag{1}$$

$$\mathfrak{H} = (H, \Gamma) \tag{2}$$

then it follows that $X = \mathfrak{G} - \Gamma$ and $Y = \mathfrak{H} - \theta$ are character groups.

Property D gives

$$\theta = (H, \mathfrak{G}) \tag{3}$$

$$\Gamma = (G, \mathfrak{H}). \tag{4}$$

Now pair the cosets $\bar{x}$ of X and $\bar{y}$ of Y by means of the pairing xy, $x \in G$, $y \in H$ which exists because G and H are character groups. If x, x' belong to $\bar{x}$, and y, y' belong to $\bar{y}$, $x - x' \in \Gamma$, $y - y' \in \theta$. Since $y \in \mathfrak{H}$, (2) gives $(x - x')y = 0$, so $xy = x'y$. Again, since $x' \subset \mathfrak{G}$, $x'(y - y') = 0$ by (1) so $x'y = x'y'$, and finally $xy = x'y'$. This makes it legitimate to define

$$\bar{x}\bar{y} = xy \quad \text{where} \quad x \in \bar{x}, y \in \bar{y}. \tag{5}$$

The two distributive laws hold for $\bar{x}\bar{y}$ since they do for xy, so X and Y are paired by (5).

Asking for those $\bar{x}$ for which $\bar{x}\bar{y} = 0$ for every $\bar{y} \in Y$ is equivalent by (5) to asking in which coset of $\mathfrak{G} - \Gamma$ x must lie in order that $xy = 0$ for every $y \in \mathfrak{H}$. Since $x \in \mathfrak{G} < G$, it follows from (4) that $x \in \Gamma$, so $\bar{x}$ is the identity of X and $(X, Y) = 0$. Similarly $(Y, X) = 0$. By B one of the pair X, Y is discrete, the other bicompact so by A they are character groups of one another.

3. Chains of the complex K

Let K be a complex made up of a finite number of elements (cells) and having its incidence relations given by matrices $M^{p,p-1} = \| u_{ij}^p \|$ whose elements are integers (including zero). The matrix $M^{p,p-1}$ determines the incidences between the p-dimensional elements (rows, first subscript) and the $(p-1)$-dimensional elements (columns, second subscript). The matrices have the property $M^{p+1,p}M^{p,p-1} = 0$.

If K has s_p p-dimensional elements, and a subcomplex L has σ_p p-dimensional elements, number these elements so that the first σ_p are in L. Then

(1) $\qquad u_{ij}^p = 0 \quad \text{when} \quad i \leqq \sigma_p \quad \text{and} \quad j > \sigma_{p-1} \quad \text{simultaneously}$

because every $(p-1)$-element incident with a p-element of L is itself in L by definition of subcomplex.

If G is an Abelian group which is either discrete or bicompact, the group of p-G-chains of K is defined to be s_pG, and that of L is σ_pG. The transformation,

$$\partial(g^1, \cdots g^{s_p}) = (f^1, \cdots, f^{s_{p-1}}),$$

$$f^j = \sum_{i=1}^{s_p} u_{ij}^p g^i \qquad j = 1, \cdots, s_{p-1},$$

is a continuous (see F) homomorphism of s_pG into $s_{p-1}G$. The subgroup ${}^p\mathfrak{Z}^G(K)$ of elements $g \in s_pG$ with $\partial g = 0$ is called the *group of p-G-cycles of K*. That ${}^p\mathfrak{Z}^G(K)$ is *closed* follows from G.

The subgroup ${}^p\mathfrak{F}^G(K)$ composed of the elements ∂g of s_pG (∂ now a transformation from $s_{p+1}G$ into s_pG) $g \in s_{p+1}G$, is called the *group of p-G-cycles bounding in K*. That ${}^p\mathfrak{F}^G(K) \ll sG$ follows from H. Notice that

$$\partial\partial g = (d^1, \cdots, d^{s_{p-1}}) \quad \text{for} \quad g \in s_{p+1}G,$$

where

$$d^k = \sum_{i=1}^{s_{p+1}} \sum_{j=1}^{s_p} u_{ij}^{p+1} u_{jk}^p g^i = 0,$$

since the product of the incidence matrices is zero. So:

THEOREM 3.1. ${}^p\mathfrak{F}^G(K) \ll {}^p\mathfrak{Z}^G(K) \ll s_pG$.

COROLLARY. ${}^p\mathfrak{F}^G(K)^{\sigma_p} \ll {}^p\mathfrak{Z}^G(L)$ (See 2c).

If g is a chain of ${}^p\mathfrak{F}^G(K)^{\sigma_p}$ it is a chain of L and $I^{-1}g$ is a chain of ${}^p\mathfrak{F}^G(K)$ so by theorem 3.1 $\partial g = 0$. The part of ∂g on L is therefore zero and $g \in {}^p\mathfrak{Z}^G(L)$.

Let H be the character group of G. Then H is either bicompact or discrete. By a p-H-chain of K is meant an element $(h_1, \cdots, h_{s_p})$ of s_pH, and similarly for L. The continuous homomorphism

$$\delta h = (k_1, \cdots k_{s_{p+1}})$$

where

$$k_i = \sum_{j=1}^{s_p} u_{ij}^{p+1} h_j \qquad i = 1, \cdots, s_{p+1}$$

carries s_pH into $s_{p+1}H$. Define ${}^p\mathfrak{Z}_H(K)$ to be the subgroup: $h \in s_pH$ and $\delta h = 0$; ${}^p\mathfrak{F}_H(K)$ to be the subgroup δh, $h \in s_{p-1}H$. The first is the *group of p-H-cocycles*; the second the *group of p-H-chains cobounding in K*. As before,

$$\delta\delta h = (t_1, \cdots, t_{s_{p+1}}) \text{ where } h \in s_{p-1}H \text{ and}$$

$$t_r = \sum_{i,j} u_{ri}^{p+1} u_{ij}^{p} h_j = 0, \text{ and:}$$

THEOREM 3.2. ${}^p\mathfrak{F}_H(K) \ll {}^p\mathfrak{Z}_H(K) \ll s_pH$.

In future s_p, σ_p will be abbreviated to s, σ.

THEOREM 3.3. ${}^p\mathfrak{F}^G(L) \ll {}^p\mathfrak{F}^G(K)^\sigma \ll \sigma G$.

a) ${}^p\mathfrak{F}^G(L) < {}^p\mathfrak{F}^G(K)^\sigma$.

If $f = (f^1, \cdots, f^\sigma) \in {}^p\mathfrak{F}^G(L)$ there are $g^1, \cdots, g^{\sigma_{p+1}}$ in G such that $f^j = \sum_{i=1}^{\sigma_{p+1}} u_{ij}^{p+1} g^i$, $j = 1, \cdots, \sigma$. Also $I^{-1}f = (f^1, \cdots, f^\sigma, 0, \cdots, 0)$. Let $g = (g^1, \cdots, g^{s_{p+1}})$ where $g^i = 0$ if $i > \sigma_{p+1}$. Then $\sum_{i=1}^{s_{p+1}} u_{ij}^{p+1} g^i = \sum_{i=1}^{\sigma_{p+1}} u_{ij}^{p+1} g^i + 0$. When $j \leqq \sigma$ the r.h.s. of this equation is f. When $j > \sigma$ the r.h.s. is zero by (1). Hence $\partial g = I^{-1}f$ and $I^{-1}f$ is in ${}^p\mathfrak{F}^G(K)$ as well as in σG, so $f \in {}^p\mathfrak{F}^G(K)^\sigma$ and a) follows.

b) Theorem 3.1 gives ${}^p\mathfrak{F}^G(L) \ll \sigma G$ and theorem 3.1 and lemma 3 give ${}^p\mathfrak{F}^G(K)^\sigma \ll \sigma G$.

COROLLARY. ${}^p\mathfrak{F}^G(L) \ll {}^p\mathfrak{F}^G(K)^\sigma \frown {}^p\mathfrak{Z}^G(L) \ll \sigma G$.

This follows from theorems 3.1 and 3.3.

THEOREM 3.4. ${}^p\mathfrak{Z}_H(K)^* \ll {}^p\mathfrak{Z}_H(L) \ll \sigma G$.

a. ${}^p\mathfrak{Z}_H(K)^* < {}^p\mathfrak{Z}_H(L)$. If $h = (h_1, \cdots, h_\sigma, \cdots, h_s)$ and $h' = (h_1, \cdots, h_\sigma)$ then $h' \in {}^p\mathfrak{Z}_H(K)^*$ implies $h \in {}^p\mathfrak{Z}_H(K)$ which gives

$$\sum_{j=1}^{s} u_{ij}^{p+1} h_j = 0, \qquad i = 1, \cdots, s_{p+1}.$$

Then

$$\sum_{j=1}^{\sigma} u_{ij}^{p+1} h_j + \sum_{j=\sigma+1}^{s} u_{ij}^{p+1} h_j = 0.$$

So

$$\textstyle\sum_{j=1}^{\sigma} u_{ij}^{p+1} h_j = 0 \qquad \text{by (1) when } i \leqq \sigma_{p+1}, \text{ and hence } h' \in {}^p\mathfrak{Z}_H(L).$$

b. Theorem 3.1 gives ${}^p\mathfrak{Z}_H(L) \ll \sigma H$ and theorem 3.2 and lemma 1 give ${}^p\mathfrak{Z}_H(K)^* \ll \sigma H$.

COROLLARY. ${}^p\mathfrak{Z}_H(K)^* \smile {}^p\mathfrak{F}_H(L) \ll {}^p\mathfrak{Z}_H(L) \ll \sigma H$.

This follows from H and theorems 3.2 and 3.4.

Two formulas from a paper by Whitney[5] are now needed to give the next two theorems.

[5] Whitney: *On Matrices of Integers and Combinatrial Topology*, Duke Journal, 3 (1937) pp. 35–45. See theorem 1 p. 36 and p. 40, (6.6) and (6.7).

THEOREM 3.5. ${}^{p}\mathfrak{F}^{G}(K) = (sG, {}^{p}\mathfrak{Z}_{H}(K))$.

THEOREM 3.6. ${}^{p}\mathfrak{F}_{H}(K) = (sH, {}^{p}\mathfrak{Z}^{G}(K))$.

THEOREM 3.7. ${}^{p}\mathfrak{F}^{G}(K)^{\sigma} = (\sigma G, \mathfrak{Z}_{H}(K)^{*})$.

By theorem 3.5, ${}^{p}\mathfrak{F}^{G}(K)^{\sigma} = (sG, {}^{p}\mathfrak{Z}_{H}(K))^{\sigma}$. But by theorem 2.1, ${}^{p}\mathfrak{F}^{G}(K)^{\sigma} = (\sigma G, {}^{p}\mathfrak{Z}_{H}(K)^{*})$.

THEOREM 3.8. ${}^{p}\mathfrak{F}^{G}(K)^{\sigma} \frown {}^{p}\mathfrak{Z}^{G}(L) = (\sigma G, {}^{p}\mathfrak{F}_{H}(L) \smile {}^{p}\mathfrak{Z}_{H}(K)^{*})$.

This is a consequence of 3.5 and E, 3.7 and 2.5.

4. Homologies and Cohomologies

If G and H are character groups, one being discrete, the other bicompact, groups involving homologies and groups involving cohomologies are paired in three theorems.

THEOREM 4.1. ${}^{p}\mathfrak{H}^{G}(K) = {}^{p}\mathfrak{Z}^{G}(K) - {}^{p}\mathfrak{F}^{G}(K)$ *and* ${}^{p}\mathfrak{H}_{H}(K) = {}^{p}\mathfrak{Z}_{H}(K) - {}^{p}\mathfrak{F}_{H}(K)$ *are character groups.*

THEOREM 4.2. ${}^{p}\mathfrak{H}^{G}(L, K) = {}^{p}\mathfrak{F}^{G}(K)^{\sigma} - {}^{p}\mathfrak{F}^{G}(L)$ *and* ${}^{p}\mathfrak{H}_{H}(L, K) = {}^{p}\mathfrak{Z}_{H}(L) - {}^{p}\mathfrak{Z}_{H}(K)^{*}$ *are character groups.*

Since ${}^{p}\mathfrak{F}^{G}(K)^{\sigma} \ll {}^{p}\mathfrak{Z}^{G}(L)$ (theorem 3.1, corollary) ${}^{p}\mathfrak{F}^{G}(K)^{\sigma} \frown {}^{p}\mathfrak{Z}^{G}(L) = {}^{p}\mathfrak{F}^{G}(K)^{\sigma}$ so ${}^{p}\mathfrak{H}^{G}(L, K) = {}^{p}\mathfrak{F}^{G}(K)^{\sigma} \frown {}^{p}\mathfrak{Z}^{G}(L) - {}^{p}\mathfrak{F}^{G}(L)$.

THEOREM 4.3. ${}^{p}\mathfrak{H}^{G}(L, K)$ *and* $Y = {}^{p}\mathfrak{Z}_{H}(L) - [{}^{p}\mathfrak{F}_{H}(L) \smile {}^{p}\mathfrak{Z}_{H}(K)^{*}]$ *are character groups.*

The proofs of these theorems are by means of B, C, D and theorem 2.6. The identifications made to obtain the notation of 2.6 are shown in the following table, as are the theorems giving conditions i and ii of 2.6.

Thm. 2.6	4.1	4.2	4.3
G	sG	σG	σG
$\mathfrak{G}$	${}^{p}\mathfrak{Z}^{G}(K)$	${}^{p}\mathfrak{F}^{G}(K)^{\sigma}$	${}^{p}\mathfrak{F}^{G}(K)^{\sigma} \frown {}^{p}\mathfrak{Z}^{G}(L)$
Γ	${}^{p}\mathfrak{F}^{G}(K)$	${}^{p}\mathfrak{F}^{G}(L)$	${}^{p}\mathfrak{F}^{G}(L)$
H	sH	σH	σH
$\mathfrak{H}$	${}^{p}\mathfrak{Z}_{H}(K)$	${}^{p}\mathfrak{Z}_{H}(L)$	${}^{p}\mathfrak{Z}_{H}(L)$
θ	${}^{p}\mathfrak{F}_{H}(K)$	${}^{p}\mathfrak{Z}_{H}(K)^{*}$	${}^{p}\mathfrak{F}_{H}(L) \smile {}^{p}\mathfrak{Z}_{H}(K)^{*}$
$X = \mathfrak{G} - \Gamma$	${}^{p}\mathfrak{H}^{G}(K)$	${}^{p}\mathfrak{H}^{G}(L, K)$	${}^{p}\mathfrak{H}^{G}(L, K)$
$Y = \mathfrak{H} - \theta$	${}^{p}\mathfrak{H}_{H}(K)$	${}^{p}\mathfrak{H}_{H}(L, K)$	Y
$\Gamma \ll \mathfrak{G} \ll G$	3.1	3.3	3.3, Corollary
$\theta \ll \mathfrak{H} \ll H$	3.2	3.4	3.4, Corollary
$\mathfrak{G} = (G, \theta)$	3.5, E	3.7	3.8
$\mathfrak{H} = (H, \Gamma)$	3.6, E	3.6, E	3.6, E

Theorem 4.1 is the well-known duality theorem for the homology and cohomology groups of a complex. Theorem 4.2 shows that *the group of p-G-cycles of L bounding in K,* mod *those of them bounding in L has for character group the group of p-H-cocycles of L,* mod *those of them which are the part on L of cocycles of K.*

THEOREM 4.4. ${}^{p}\mathfrak{H}_{H}(L, K) \approx {}^{p}\mathfrak{H}_{H}(L) - {}^{p}\Delta_{H}(L, K)$, *where* ${}^{p}\Delta_{H}(L, K)$ *is the group of cosets of* ${}^{p}\mathfrak{H}_{H}(L)$ *each of which contains the part on L of a cocycle of K.*

By theorem 4.3, Y being the character group of ${}^{p}\mathfrak{H}^{G}(L, K)$ is isomorphic to ${}^{p}\mathfrak{H}_{H}(L, K)$. Then by theorem 2.4 $Y = {}^{p}\mathfrak{H}_{H}(L) - {}^{p}\Delta_{H}(L, K)$ where ${}^{p}\Delta_{H}(L, K)$ is the B^{A} of that theorem.

CORNELL UNIVERSITY.

ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

ON THE THEORY OF STATIONARY RANDOM PROCESSES

BY HARALD CRAMÉR

(Received May 5, 1939)

I. INTRODUCTION

1. Let t be a real parameter which may be regarded as representing time and let, for every t, $X(t)$ denote a real-valued random variable.

Following Khintchine [6], we shall say that $X(t)$ is attached to a *one-dimensional random process* if, for any finite set of real values $t_1, \cdots, t_k$, the simultaneous probability distribution[1] of the variables $X(t_1), \cdots, X(t_k)$ is known. This distribution may, *e.g.*, be characterized by means of the corresponding k-dimensional distribution function $F(x_1, \cdots, x_k; t_1, \cdots, t_k)$, which represents the probability of the set of inequalities

$$X(t_\nu) \leqq x_\nu, \qquad (\nu = 1, 2, \cdots, k).$$

F is a distribution function in the variables $x_1, \cdots, x_k$, with $t_1, \cdots, t_k$ as parameters.

Any system of distribution functions $F(x_1, \cdots, x_k; t_1, \cdots, t_k)$, where $k = 1, 2, \cdots$, such that every F is symmetric in all pairs (x_ν, t_ν) and satisfies the *consistency relation*

$$F(x_1, \cdots, x_j, \infty, \cdots, \infty; t_1, \cdots, t_k) = F(x_1, \cdots, x_j; t_1, \cdots, t_j)$$

for any $j < k$, defines a random process.

For certain purposes this definition is not complete enough, and Doob [3] has given more elaborate definitions which meet the requirements of these cases. In the present paper we shall, however, only deal with problems where the simple definition given above is sufficient.

The definition is immediately extended to an *n-dimensional process*, if an n-dimensional random variable $Z(t) = \{X_1(t), \cdots, X_n(t)\}$ is chosen as the basic variable of the process. The simultaneous distribution of the variables $Z(t_1), \cdots, Z(t_k)$ is then characterized by a distribution function in a space of nk dimensions, which represents the probability of the set of inequalities

$$X_\mu(t_\nu) \leqq x_{\mu\nu}, \qquad (\mu = 1, 2, \cdots, n; \nu = 1, 2, \cdots, k).$$

The corresponding extension of the conditions of symmetry and consistency is obvious.

---

[1] Cf the writer's Cambridge Tract [1] on "Random Variables and Probability Distributions."

Finally, we may evidently assume the n components $X_\mu(t)$ of the basic variable $Z(t)$ to be *complex random variables*

$$X_\mu(t) = U_\mu(t) + i\, V_\mu(t), \tag{1}$$

and so obtain a *complex multi-dimensional process*. The simultaneous distribution of $Z(t_1), \cdots, Z(t_k)$ is then characterized by a distribution function F in a space of $2nk$ dimensions, which represents the probability of the set of inequalities

$$U_\mu(t_\nu) \leqq u_{\mu\nu}\,,$$

$$V_\mu(t_\nu) \leqq v_{\mu\nu}\,,$$

where $\mu = 1, 2, \cdots, n$; $\nu = 1, 2, \cdots, k$. F is a distribution function in the arguments $u_{\mu\nu}$ and $v_{\mu\nu}$, with $t_1, \cdots, t_k$ as parameters.

2. Consider a random process defined by

$$Z(t) = \{X_1(t), \cdots, X_n(t)\}$$

where, in the general case, the $X_\mu(t)$ are complex random variables defined according to (1). It evidently amounts to the same thing if we consider *simultaneously* the n processes defined by the components $X_1(t), \cdots, X_n(t)$, and we shall apply these two points of view without distinction in the sequel.

We shall always assume that the mean values $E\{|\, X_\mu(t)\,|^2\}$ are finite for all μ and t. It then follows that the mean values $E\{X_\mu(t)\}$ and $E\{X_\mu(t)\overline{X_\nu(u)}\}$ are finite for all μ, ν, t and u.

Generalizing a definition introduced by Khintchine [6][2], we shall say that $Z(t)$ defines a *stationary random process*, if the following two conditions are satisfied for all μ and ν:

(A) $\quad E\{X_\mu(t)\} = m_\mu$ *is independent of* t;

(B) $\quad E\{X_\mu(t)\overline{X_\nu(u)}\} = R_{\mu\nu}(t-u)$ *is a function of the difference* $t-u$.

If these conditions are satisfied, we shall also say that the n processes defined by the variables $X_1(t), \cdots, X_n(t)$ are *stationary* and *stationarily correlated*.

Obviously we may without restriction of generality assume

$$\text{(C)} \qquad \begin{cases} E\{X_\mu(t)\} = m_\mu = 0, \\ E\{|\, X_\mu(t)\,|^2\} = R_{\mu\mu}(0) = \sigma_\mu^2 > 0 \end{cases}$$

for all μ. This being assumed, we shall denote the $R_{\mu\nu}(t)$ as the *correlation functions* of the stationary processes attached to the variables $X_1(t), \cdots, X_n(t)$. More particularly, we shall call $R_{\mu\mu}(t)$ the *autocorrelation function* of the process defined by $X_\mu(t)$, while $R_{\mu\nu}(t)$ and $R_{\nu\mu}(t)$ will, for $\mu \neq \nu$, be called the *mutual*

[2] Khintchine introduces also another, narrower, concept of stationarity which, however, we shall have no occasion to use in the present paper.

correlation functions of the processes connected with $X_\mu(t)$ and $X_\nu(t)$. It follows from (B) that we always have

$$R_{\nu\mu}(-t) = \overline{R_{\mu\nu}(t)}. \tag{2}$$

In the particular case when the variables $X_\mu(t)$ are real, $R_{\mu\nu}(t)/(\sigma_\mu\sigma_\nu)$ is evidently identical with the ordinary correlation coefficient of $X_\mu(t_0 + t)$ and $X_\nu(t_0)$.

We shall finally assume that

$$\lim_{t\to 0} R_{\mu\mu}(t) = R_{\mu\mu}(0) = \sigma_\mu^2 \tag{D}$$

for $\mu = 1, 2, \cdots, n$. It will be shown later (Lemma 1) that, if this condition is satisfied, all the correlation functions $R_{\mu\nu}(t)$ are continuous for all real values of t. The process defined by each $X_\mu(t)$, as well as the one defined by the composite variable $Z(t)$, will then be called a *continuous stationary random process.* This concept has been introduced by Khintchine [6], who has justly emphasized its great importance for various branches of applications.

3. The main purpose of the present paper is to study the mutual relations of the correlation functions $R_{\mu\nu}(t)$. In Chapter III, we shall deduce a necessary and sufficient condition that a set of n^2 *a priori* given functions $R_{\mu\nu}(t)$ may be correlation functions of n complex random processes satisfying the conditions (A)–(D). It will be shown to be necessary and sufficient that the $R_{\mu\nu}(t)$ should, for all real t, be given by Fourier-Stieltjes integrals of the form

$$R_{\mu\nu}(t) = \int_{-\infty}^{\infty} e^{itx}\, dF_{\mu\nu}(x),$$

where the $F_{\mu\nu}(x)$ are functions of bounded variation in $(-\infty, \infty)$ which satisfy certain conditions as specified by Theorem 1. In particular, it follows from these conditions that the functions $F_{\mu\mu}(x)$ corresponding to the autocorrelation functions $R_{\mu\mu}(t)$ are real and never decreasing, with a total variation equal to σ_μ^2, as already proved by Khintchine [6].

In Chapter IV, some further properties of the correlation functions will be deduced, and the particular case of random processes defined by *real* variables $X_1(t)$ and $X_2(t)$ will be considered. Finally, in Chapter V we shall show that an analogous theory may be developed for the case when our random processes are considered for *integral values* only of the parameter t.

II. Auxiliary Theorems

4. In this Chapter we shall, for the sake of convenience, give as Lemmas a number of results (partly previously known) wanted in the sequel. In the first place, we shall prove the following result indicated in the Introduction.

Lemma 1. *Let $R_{\mu\nu}(t)$ be the n^2 correlation functions of n random processes satisfying the conditions* (A)–(D) *of paragraph* 2. *Then $R_{\mu\nu}(t)$ is continuous for all real t, and we have*

$$|R_{\mu\nu}(t)| \leqq \sigma_\mu\sigma_\nu. \tag{3}$$

The inequality (3) follows immediately from (B) and (C) by means of the Schwarz inequality. Further, we have

$$R_{\mu\nu}(t + \Delta t) - R_{\mu\nu}(t) = E\{\overline{X_\nu(0)}[X_\mu(t + \Delta t) - X_\mu(t)]\}$$

and thus by the Schwarz inequality

$$\begin{aligned} |R_{\mu\nu}(t + \Delta t) - R_{\mu\nu}(t)| &\leqq \sigma_\nu \sqrt{E\{|X_\mu(t + \Delta t) - X_\mu(t)|^2\}} \\ &= \sigma_\nu \sqrt{2\sigma_\mu^2 - R_{\mu\mu}(\Delta t) - R_{\mu\mu}(-\Delta t)}. \end{aligned}$$

According to (D), this tends to zero with Δt.

5. We next consider some convergence properties of sequences of random variables. If $X_1, X_2, \cdots$ and X are real or complex random variables, we shall say that X_n *converges in the mean* to X, and write

$$\underset{n\to\infty}{\text{l.i.m.}}\ X_n = X,$$

whenever the mean value

$$E\{|X_n - X|^2\}$$

is finite for all sufficiently large n and tends to zero as $n \to \infty$.

If the mean value

$$E\{|X_m - X_n|^2\}$$

tends to zero as m and n tend independently to infinity, then it is known[3] that there is a variable X such that X_n converges in the mean to X. If, in addition, $E\{|X_n|^2\}$ is finite for all sufficiently large n, it immediately follows that $E\{|X|^2\}$ is finite. Further, if X' and X'' are variables such that X_n converges in the mean both to X' and to X'', then X' and X'' are equivalent, i.e.

$$Pr\{X' = X''\} = 1.$$

LEMMA 2. *Let $X_1, X_2, \cdots$ and $Y_1, Y_2, \cdots$ be two sequences of random variables such that $E\{|X_n|^2\}$ and $E\{|Y_n|^2\}$ are finite for all n. Suppose that the mean values $E\{|X_m - X_n|^2\}$ and $E\{|Y_m - Y_n|^2\}$ both tend to zero as m and n tend independently to infinity, so that there are two variables X and Y such that $E\{|X|^2\}$ and $E\{|Y|^2\}$ are finite, and* l.i.m. $X_n = X$, l.i.m. $Y_n = Y$. *Then*

$$\begin{aligned} &E(X_n) \to E(X), && E(Y_n) \to E(Y), \\ &E\{|X_n|^2\} \to E\{|X|^2\}, && E\{|Y_n|^2\} \to E\{|Y|^2\}, \\ &E(X_n Y_n) \to E(XY). \end{aligned}$$

---

[3] Cf. e.g. Fréchet [4], Lévy [7].—It is also known that convergence in the mean implies convergence "in probability," which is expressed by the relation $Pr\{|X_n - X| > \epsilon\} \to 0$ for any $\epsilon > 0$.

In fact, we obtain by means of the Schwarz inequality

$$|E(X_n) - E(X)|^2 \leqq E^2\{|X_n - X|\} \leqq E\{|X_n - X|^2\},$$

and

$$\begin{aligned}[E\{|X_n|^2\} - E\{|X|^2\}]^2 &= E^2\{(|X_n| - |X|)(|X_n| + |X|)\} \\ &\leqq E\{(|X_n| - |X|)^2\} \cdot E\{(|X_n| + |X|)^2\} \\ &\leqq E\{|X_n - X|^2\} \cdot E\{(|X_n - X| + 2|X|)^2\} \\ &\leqq 2E\{|X_n - X|^2\} \cdot E\{|X_n - X|^2 + 4|X|^2\}\end{aligned}$$

Hence we immediately conclude that $E(X_n) \to E(X)$ and $E\{|X_n|^2\} \to E\{|X|^2\}$. Applying the last result to the sequences $X_n + \bar{Y}_n$ and $X_n + i\bar{Y}_n$, we find that $E(X_n Y_n) \to E(XY)$.

6. In the sequel, we shall repeatedly have to deal with integrals containing a random variable attached to a continuous stationary random process. Integrals of this kind may be very simply defined in complete analogy to the elementary definition of an ordinary Riemann integral, as shown by the following Lemma.

LEMMA 3. *Let $X(t)$ be a (complex) random variable, attached to a continuous stationary random process, and let $\varphi(t)$ be a continuous function of t for $a \leqq t \leqq b$. Consider a sequence of random variables $Z_1, Z_2, \cdots$, defined by*

$$Z_n = \sum_{\nu=1}^{n} \varphi(t_\nu^{(n)}) X(t_\nu^{(n)})(t_\nu^{(n)} - t_{\nu-1}^{(n)}),$$

where $a = t_0^{(n)} < t_1^{(n)} < \cdots < t_{n-1}^{(n)} < t_n^{(n)} = b$. Suppose that

$$\operatorname*{Max}_{\nu=1,2,\cdots,n} (t_\nu^{(n)} - t_{\nu-1}^{(n)}) \to 0$$

as $n \to \infty$. Then Z_n converges in the mean to a random variable Z. Further, if $\{Z_n^*\}$ is another sequence formed with the same $X(t)$ and $\varphi(t)$, but with another system of points $t_\nu^{(n)}$ satisfying the above conditions, and converging in the mean to the limit Z^*, then Z^* and Z are equivalent. For the limit Z we write, by definition,

$$Z = \operatorname*{l.i.m.}_{n\to\infty} Z_n = \int_a^b \varphi(t) X(t)\, dt,$$

so that the last integral is a random variable, which is hereby completely defined but for equivalence.

It is evident that we may assume $E\{X(t)\} = 0$.

Putting

$$Z' = \sum_{\nu=1}^{m} \varphi(t_\nu') X(t_\nu')(t_\nu' - t_{\nu-1}'),$$

$$Z'' = \sum_{\nu=1}^{n} \varphi(t_\nu'') X(t_\nu'')(t_\nu'' - t_{\nu-1}''),$$

it will obviously only be necessary to show that the mean value $E\{|Z' - Z''|^2\}$ can be made as small as we please by taking the length δ of the maximum interval in the two divisions corresponding to the points t_ν' and t_ν'' sufficiently small. In fact, this point being established, it directly follows that any sequence $\{Z_n\}$ or $\{Z_n^*\}$ converges in the mean, and that the mean value $E\{|Z_n - Z_n^*|^2\}$ tends to zero. Hence the limits Z and Z^* must be equivalent.

From the current text-book theory of the ordinary Riemann integral, it is well known that we may restrict ourselves to the case when the division points t_ν' ($\nu = 0, 1, \cdots, m$) *are all contained among the points* t_ν'' ($\nu = 0, 1, \cdots, n$). This being assumed, we may write

$$Z' - Z'' = \sum_{\nu=1}^{n} \{\varphi(\tau_\nu)X(\tau_\nu) - \varphi(t_\nu'')X(t_\nu'')\}(t_\nu'' - t_{\nu-1}'')$$

where

$$|\tau_\nu - t_\nu''| < \delta.$$

Denoting by $R(t)$ the autocorrelation function of our continuous stationary process, we then have

$$E\{|Z' - Z''|^2\}$$

$$= \sum_{\mu,\nu} \varphi(\tau_\mu)\{\overline{\varphi(\tau_\nu)}R(\tau_\mu - \tau_\nu) - \overline{\varphi(t_\nu'')}R(\tau_\mu - t_\nu'')\}(t_\mu'' - t_{\mu-1}'')(t_\nu'' - t_{\nu-1}'')$$

$$- \sum_{\mu,\nu} \varphi(t_\mu'')\{\overline{\varphi(\tau_\nu)}R(t_\mu'' - \tau_\nu) - \overline{\varphi(t_\nu'')}R(t_\mu'' - t_\nu'')\}(t_\mu'' - t_{\mu-1}'')(t_\nu'' - t_{\nu-1}'').$$

$\varphi(t)$ and $R(t)$ being both continuous functions, it easily follows that the last expression can be made less than any given ϵ by taking δ sufficiently small. Thus the Lemma is proved.

It is easily shown that the elementary properties of the ordinary Riemann integral hold for the integral of a random variable as defined by Lemma 3. Integrals of this type are themselves random variables, and we shall have occasion to consider in the sequel certain mean values connected with such variables. In this respect, we shall require the following Lemma.

LEMMA 4. *Let $X_1(t)$ and $X_2(t)$ be random variables attached to continuous stationary processes satisfying the conditions* (A)–(D) *of paragraph* 2, *and let $\varphi_1(t)$ and $\varphi_2(t)$ be continuous functions of t for $a \leqq t \leqq b$. Then*

$$(4) \qquad E\left\{\int_a^b \varphi_1(t)X_1(t)\,dt \cdot \int_a^b \varphi_2(u)\overline{X_2(u)}\,du\right\} = \int_a^b \int_a^b \varphi_1(t)\varphi_2(u)R_{12}(t-u)\,dt\,du.$$

In particular, for $X_1(t) = X_2(t)$, $\varphi_1(t) = \overline{\varphi_2(t)}$, we have

$$E\left\{\left|\int_a^b \varphi_1(t)X_1(t)\,dt\right|^2\right\} = \int_a^b \int_a^b \varphi_1(t)\overline{\varphi_1(u)}R_{11}(t-u)\,dt\,du.$$

The last expression, being the mean value of a real and non-negative variable, is evidently real and non-negative.

In order to prove this Lemma, we have only to express the integrals in the first member of (4) as limits of finite sums of random variables according to Lemma 3. The mean value of the expression thus obtained is equal to a Riemann sum corresponding to the double integral in the second member of (4). An application of the last relation of Lemma 2 then immediately yields the desired result.

III. Main Theorem on the Correlation Functions

7. We now proceed to the proof of the following theorem, which constitutes the main result of the present paper.[4]

THEOREM 1. *a) Let $X_1(t), \cdots, X_n(t)$ be complex random variables attached to n continuous stationary processes satisfying the conditions* (A)–(D) *of paragraph* 2. *Then the correlation functions defined by* (B) *are for all real t given by Fourier-Stieltjes integrals of the form*

$$R_{\mu\nu}(t) = \int_{-\infty}^{\infty} e^{itx}\, dF_{\mu\nu}(x), \tag{5}$$

where the $F_{\mu\nu}(x)$ are functions of bounded variation in $(-\infty, \infty)$, which we may always assume to be everywhere continuous to the right. Further, denoting by $\Delta F_{\mu\nu}$ the increase $F_{\mu\nu}(\beta) - F_{\mu\nu}(\alpha)$ of $F_{\mu\nu}(x)$ over the real interval $\alpha \leqq x \leqq \beta$, the form

$$H(z_1, \cdots, z_n) = \sum_{\mu,\nu=1}^{n} z_\mu \bar{z}_\nu \Delta F_{\mu\nu} \tag{6}$$

is for any closed interval (α, β) a non-negative Hermite form. (In particular, putting all the z-variables except z_μ equal to zero, it follows that $F_{\mu\mu}(x)$ is real and never decreasing.)

b) Conversely, let $R_{\mu\nu}(t)$ be a set of n^2 functions given by the Fourier-Stieltjes integrals (5), *the functions $F_{\mu\nu}(x)$ being of bounded variation in $(-\infty, \infty)$ and everywhere continuous to the right, and suppose that the form* (6) *is, for any closed real interval (α, β), a non-negative Hermite form. Then it is possible to find n continuous stationary random processes having the correlation functions $R_{\mu\nu}(t)$.*

8. In order to prove the first part of the theorem, we shall begin by deducing the expression (5) for the correlation functions. We put, for any real x and any $A > 1$,

$$g_{\mu\nu}(x, A) = \frac{1}{2\pi A} \int_0^A \int_0^A R_{\mu\nu}(t - u) e^{-ix(t-u)}\, dt\, du. \tag{7}$$

---

[4] Professor A. Kolmogoroff has kindly informed the author that the same theorem has been found by him.

According to Lemma 1, $R_{\mu\nu}(t)$ is a bounded and continuous function, and it is then known[5] that a necessary and sufficient condition for a representation of $R_{\mu\nu}(t)$ of the form (5), where $F_{\mu\nu}(x)$ is of bounded variation in $(-\infty, \infty)$, is

$$\int_{-\infty}^{\infty} |g_{\mu\nu}(x, A)|\, dx < \text{const.} \tag{8}$$

for all $A > 1$. Moreover, a necessary and sufficient condition for a representation of the form (5) with a real and never decreasing $F_{\mu\nu}(x)$ is that $g_{\mu\nu}(x, A)$ should be real and non-negative for all real x and all $A > 1$. If the latter condition is satisfied, then (8) holds *a fortiori*.

Now according to Lemma 4 we have

$$\begin{aligned} &\frac{1}{2\pi A} E\left\{\int_0^A X_\mu(t)e^{-ixt}\, dt \cdot \int_0^A \overline{X_\nu(u)}e^{ixu}\, du\right\} \\ &\qquad = \frac{1}{2\pi A}\int_0^A \int_0^A R_{\mu\nu}(t-u)e^{-ix(t-u)}\, dt\, du = g_{\mu\nu}(x, A). \end{aligned} \tag{9}$$

Thus in particular we have for $\mu = \nu$

$$g_{\mu\mu}(x, A) = \frac{1}{2\pi A} E\left\{\left|\int_0^A X_\mu(t)e^{-ixt}\, dt\right|^2\right\},$$

which shows that $g_{\mu\mu}(x, A)$ is real and non-negative, so that we have

$$R_{\mu\mu}(t) = \int_{-\infty}^{\infty} e^{ixt}\, dF_{\mu\mu}(x),$$

where $F_{\mu\mu}(x)$ is real and never decreasing. From $|R_{\mu\mu}(t)| \leqq \sigma_\mu^2$ it then follows (Cramér [2], formula (6)) that

$$\int_{-\infty}^{\infty} g_{\mu\mu}(x, A)\, dx \leqq \sigma_\mu^2. \tag{10}$$

On the other hand, we obtain from (9), using the Schwarz inequality,

$$\begin{aligned} |g_{\mu\nu}(x, A)|^2 &= \frac{1}{(2\pi A)^2}\left|E\left\{\int_0^A X_\mu(t)e^{-ixt}\, dt \cdot \int_0^A \overline{X_\nu(u)}e^{ixu}\, du\right\}\right|^2 \\ &\leqq \frac{1}{2\pi A} E\left\{\left|\int_0^A X_\mu(t)e^{-ixt}\, dt\right|^2\right\} \cdot \frac{1}{2\pi A} E\left\{\left|\int_0^A X_\nu(t)e^{-ixt}\, dt\right|^2\right\} \\ &= g_{\mu\mu}(x, A) \cdot g_{\nu\nu}(x, A). \end{aligned}$$

Hence we deduce by (10), using once more the Schwarz inequality,

$$\int_{-\infty}^{\infty} |g_{\mu\nu}(x, A)|\, dx \leqq \sigma_\mu \sigma_\nu.$$

---

[5] Cramér [2].

Thus the condition (8) is satisfied, and the representation (5) is established. Obviously, we may always assume that the functions $F_{\mu\nu}(x)$ are everywhere continuous to the right, since a modification of $F_{\mu\nu}(x)$ in an enumerable set of points does not affect the value of the integral in the second member of (5).

It remains to prove the assertion concerning the form $H(z_1, \cdots, z_n)$ defined by (6). Obviously it is only necessary to show that, for arbitrary complex $z_1, \cdots, z_n$, this form is always real and non-negative.

Let (α, β) be a given real interval and put

$$\varphi(t) = \frac{e^{-it\alpha} - e^{-it\beta}}{2\pi it}.$$

According to Lemma 4, the mean value

$$H_M(z_1, \cdots, z_n) = E\left\{\left|\int_{-M}^{M} \varphi(t) \sum_{\mu=1}^{n} z_\mu X_\mu(t)\, dt\right|^2\right\}$$

exists for any $M > 0$ and for arbitrary complex z_μ, and is always real and non-negative. We have, however, by Lemma 4

$$\begin{aligned} H_M(z_1, \cdots, z_n) &= \sum_{\mu,\nu} z_\mu \bar{z}_\nu E\left\{\int_{-M}^{M} \varphi(t) X_\mu(t)\, dt \cdot \int_{-M}^{M} \overline{\varphi(u) X_\nu(u)}\, du\right\} \\ &= \sum_{\mu,\nu} z_\mu \bar{z}_\nu \int_{-M}^{M}\int_{-M}^{M} \varphi(t)\overline{\varphi(u)} R_{\mu\nu}(t-u)\, dt\, du \qquad (11) \\ &= \sum_{\mu,\nu} z_\mu \bar{z}_\nu \int_{-\infty}^{\infty} dF_{\mu\nu}(x) \left|\int_{-M}^{M} \varphi(t) e^{itx}\, dt\right|^2. \end{aligned}$$

Let us first suppose that α and β have been chosen such that the functions $F_{\mu\nu}(x)$ are all continuous for $x = \alpha$ and for $x = \beta$. Given any $\epsilon > 0$, the expression

$$\int_{-M}^{M} \varphi(t) e^{itx}\, dt = \frac{1}{\pi}\int_0^M \frac{\sin t(x-\alpha)}{t}\, dt - \frac{1}{\pi}\int_0^M \frac{\sin t(x-\beta)}{t}\, dt$$

tends for indefinitely increasing M to the limit zero, uniformly for $x < \alpha - \epsilon$ and for $x > \beta + \epsilon$, and to the limit unity, uniformly for $\alpha + \epsilon < x < \beta - \epsilon$. In the remaining intervals $|x - \alpha| \leqq \epsilon$ and $|x - \beta| \leqq \epsilon$, we have

$$\left|\int_{-M}^{M} \varphi(t) e^{itx}\, dt\right| < 2.$$

It then follows from (6) and (11) that we have

$$\lim_{M\to\infty} H_M(z_1, \cdots, z_n) = H(z_1, \cdots, z_n).$$

$H_M(z_1, \cdots, z_n)$ being, according to the above, real and non-negative, the same holds true for $H(z_1, \cdots, z_n)$. As the functions $F_{\mu\nu}(x)$ are everywhere continuous to the right, this result is immediately extended to arbitrary real values of α and β, and the first part of the theorem is proved.

9. Let $R_{\mu\nu}(t)$ be a set of n^2 given functions satisfying the conditions of the second part of the theorem. We then have to show that we can find n continuous stationary random processes having the functions $R_{\mu\nu}(t)$ for their correlation functions. It will, in fact, be shown that n processes with the given correlation functions can be defined by random variables of the form

$$X_\mu(t) = U_\mu(t) + iV_\mu(t), \qquad (\mu = 1, 2, \cdots, n) \tag{12}$$

in such a way that the simultaneous probability distribution of the $2nk$ real random variables

$$U_\mu(t_r), \qquad V_\mu(t_r), \qquad (\mu = 1, 2, \cdots, n; r = 1, 2, \cdots, k), \tag{13}$$

is always a *normal* distribution.

A normal distribution in any number of variables[6] is completely defined by its first and second order moments. The functions $R_{\mu\nu}(t)$ being given, we now define the normal distribution of the $2nk$ variables (13) by putting

$$\begin{cases} E\{U_\mu(t_r)\} = E\{V_\mu(t_r)\} = 0, \\ E\{U_\mu(t_r)U_\nu(t_s)\} = E\{V_\mu(t_r)V_\nu(t_s)\} = \dfrac{R_{\mu\nu}(t_r - t_s) + R_{\nu\mu}(t_s - t_r)}{4}, \\ E\{U_\mu(t_r)V_\nu(t_s)\} = \dfrac{-R_{\mu\nu}(t_r - t_s) + R_{\nu\mu}(t_s - t_r)}{4i}. \end{cases} \tag{14}$$

From our hypotheses concerning the given functions $R_{\mu\nu}(t)$, it easily follows that the relation (2) is satisfied, so that the second order moments defined by (14) are real. In order to show that these moments really define a normal distribution in the variables (13), it is sufficient to show[7] that the quadratic form $Q(u, v)$ in the $2nk$ real variables $u_{\mu r}$, $v_{\mu r}$, defined by

$$\begin{aligned} Q(u, v) &= E\{[\sum_{\mu,r} u_{\mu r} U_\mu(t_r) + v_{\mu r} V_\mu(t_r)]^2\} \\ &= \sum_{\mu,r,\nu,s} \left\{ (u_{\mu r} u_{\nu s} + v_{\mu r} v_{\nu s}) \frac{R_{\mu\nu}(t_r - t_s) + R_{\nu\mu}(t_s - t_r)}{4} \right. \\ &\qquad \left. + (u_{\mu r} v_{\nu s} - u_{\nu s} v_{\mu r}) \frac{-R_{\mu\nu}(t_r - t_s) + R_{\nu\mu}(t_s - t_r)}{4i} \right\} \end{aligned}$$

is non-negative. If this is so, the moments (14) define a normal distribution with the characteristic function

$$e^{-\frac{1}{2}Q(u,v)}.$$

Using (2) we obtain, however, by some easy transformations

$$Q(u, v) = \tfrac{1}{2} \sum_{\mu,r,\nu,s} (u_{\mu r} - iv_{\mu r})(u_{\nu s} + iv_{\nu s}) R_{\mu\nu}(t_r - t_s).$$

[6] Cf. Cramér [1], p. 109.

[7] Cf. Cramér, l.c.

Substituting here the expression (5) for $R_{\mu\nu}(t)$, we obtain further

$$Q(u, v) = \tfrac{1}{2} \int_{-\infty}^{\infty} \sum_{\mu,\nu} z_\mu \bar{z}_\nu \, dF_{\mu\nu}(x),$$

where

$$z_\mu = z_\mu(x) = \sum_r (u_{\mu r} - iv_{\mu r}) e^{it_r x}.$$

It now follows from our hypothesis concerning the form (6) that $Q(u, v)$ is never negative.

Thus the moments (14) define a normal distribution in the $2nk$ variables (13). Evidently this distribution satisfies the necessary symmetry and consistence conditions, so that the corresponding complex variables (12) define a multi-dimensional random process. From (14) we obtain, however,

$$\begin{aligned} E\{X_\mu(t_r)\overline{X_\nu(t_s)}\} &= E\{U_\mu(t_r)U_\nu(t_s) + V_\mu(t_r)V_\nu(t_s) - iU_\mu(t_r)V_\nu(t_s) + iU_\nu(t_s)V_\mu(t_r)\} \\ &= R_{\mu\nu}(t_r - t_s). \end{aligned}$$

Since t_r and t_s are arbitrary, this shows that the n random processes defined by the variables (12) are stationary and stationarily correlated processes, the correlation functions of which are identical with the given functions $R_{\mu\nu}(t)$. Thus our theorem is completely proved.

IV. Further Properties of the Correlation Functions. The Case of Real Processes

10. Let us first observe that the form H defined by (6) is non-negative and Hermitian for any interval (α, β) if, and only if, the same holds true when $\Delta F_{\mu\nu}$ denotes the total increase

$$\sum_r [F_{\mu\nu}(\beta_r) - F_{\mu\nu}(\alpha_r)]$$

of $F_{\mu\nu}(x)$ over any finite or enumerable set of non-overlapping closed intervals (α_r, β_r). Thus Theorem 1 remains valid if we introduce this more general definition of $\Delta F_{\mu\nu}$.

Without restriction of generality, we may assume $F_{\mu\nu}(-\infty) = 0$. Any $F_{\mu\nu}(x)$ being of bounded variation in $(-\infty, \infty)$ and everywhere continuous to the right, it is known that its points of discontinuity x_r form at most an enumerable set, that a finite derivative $F'_{\mu\nu}(x)$ exists almost everywhere, and that we may write for all x

$$F_{\mu\nu}(x) = F^{(a)}_{\mu\nu}(x) + F^{(d)}_{\mu\nu}(x) + F^{(s)}_{\mu\nu}(x). \tag{15}$$

Here

$$F^{(a)}_{\mu\nu}(x) = \int_{-\infty}^{x} F'_{\mu\nu}(t)\, dt$$

(the integral being taken in the Lebesgue sense) is the *absolutely continuous component* of $F_{\mu\nu}$;

$$F_{\mu\nu}^{(d)}(x) = \sum_{x_r \leqq x} p_r$$

is the *discontinuous component* (p_r being the saltus of $F_{\mu\nu}$ at $x = x_r$); finally $F_{\mu\nu}^{(s)}(x)$ is the *singular component.* $F_{\mu\nu}^{(s)}(x)$ is everywhere continuous and has almost everywhere a derivative equal to zero. There is a set Z of measure zero such that the total variation of $F_{\mu\nu}^{(s)}(x)$ over Z is equal to the total variation of $F_{\mu\nu}^{(s)}(x)$ over $(-\infty, \infty)$. The three components in the representation of $F_{\mu\nu}(x)$ according to (15) are uniquely determined by $F_{\mu\nu}(x)$.

Substituting the expression (15) for $F_{\mu\nu}(x)$ in the definition (6) of the form $H(z_1, \cdots, z_n)$, we obtain $H = H_a + H_d + H_s$, where

$$H_a(z_1, \cdots, z_n) = \sum_{\mu,\nu=1}^{n} z_\mu \bar{z}_\nu \Delta F_{\mu\nu}^{(a)}$$

and similarly for H_d and H_s. We then have the following theorem.

THEOREM 2. *If the form H is, for any interval (α, β), a non-negative Hermite form, then the same holds true for each of the forms H_a, H_d and H_s. The converse is obvious.*

Denoting by x any point such that all derivatives $F'_{\mu\nu}(x)$ exist, and observing that, by hypothesis, the form H corresponding to the interval $(x, x + dx)$ is non-negative, we find that

$$\sum_{\mu,\nu=1}^{n} z_\mu \bar{z}_\nu F'_{\mu\nu}(x)$$

is, for almost all x, a non-negative Hermite form. Integrating between the limits α and β, it follows that the same holds true for $H_a(z_1, \cdots, z_n)$.

Similarly we find, denoting by $p_{\mu\nu}$ the saltus of $F_{\mu\nu}$ in a certain point x and considering the form H corresponding to the interval $(x - \epsilon, x + \epsilon)$, that the form $\sum_{\mu,\nu=1}^{n} z_\mu \bar{z}_\nu p_{\mu\nu}$ is a non-negative Hermite form. Adding together the analogous forms for all discontinuities x belonging to (α, β), we obtain the desired result for H_d.

Finally, it is seen without difficulty that, to any closed interval (α, β) and any given $\epsilon > 0$, we can find a sub-set E of (α, β) consisting of a finite or enumerable number of non-overlapping closed intervals of total length $< \epsilon$, such that the total variation of any $F_{\mu\nu}^{(s)}$ over E differs by less than ϵ both from the total variation of $F_{\mu\nu}^{(s)}$ over (α, β) and from the total variation of $F_{\mu\nu}$ over E. Denoting by $\Delta F_{\mu\nu}$ the increase of $F_{\mu\nu}$ over E, the corresponding form H is, according to the above, a non-negative Hermite form. Allowing now ϵ to tend to zero, we obtain the form H_s corresponding to the interval (α, β), and the proof of the theorem is completed.

11. Consider now the particular case $n = 2$ of our previous theorems. We then have two continuous stationary processes, defined by the random variables

$X_1(t)$ and $X_2(t)$. For $n = 2$, the form H defined by (6) is a non-negative Hermite form if, and only if, F_{11} and F_{22} are real and never decreasing, $\Delta F_{12} = \overline{\Delta F}_{21}$, and

$$|\Delta F_{12}|^2 \leqq \Delta F_{11} \cdot \Delta F_{22}. \tag{16}$$

From Theorems 1 and 2, we then immediately obtain the following result.

THEOREM 3. *Four given functions $R_{\mu\nu}(t)$, $(\mu, \nu = 1, 2)$, are the correlation functions of two continuous stationary processes if, and only if, the $R_{\mu\nu}(t)$ are given by the Fourier-Stieltjes integrals* (5), *where F_{11} and F_{22} are real and never decreasing, $F_{21} = \bar{F}_{12}$, and*

$$|\Delta F_{12}^{(a)}|^2 \leqq \Delta F_{11}^{(a)} \cdot \Delta F_{22}^{(a)}, \tag{17a}$$

$$|\Delta F_{12}^{(d)}|^2 \leqq \Delta F_{11}^{(d)} \cdot \Delta F_{22}^{(d)}, \tag{17d}$$

$$|\Delta F_{12}^{(s)}|^2 \leqq \Delta F_{11}^{(s)} \cdot \Delta F_{22}^{(s)}, \tag{17s}$$

for any interval (α, β), the notation being the same as that used in connection with (15).

The condition (17a) is obviously equivalent to the condition that almost everywhere

$$|F_{12}'|^2 \leqq F_{11}' F_{22}'.$$

Further, (17d) is equivalent to the condition that the discontinuity points of F_{12} form a sub-set of the set of points which are discontinuity points both of F_{11} and of F_{22}, and that the saltuses $p_{\mu\nu}$ corresponding to any discontinuity point satisfy the inequality

$$|p_{12}|^2 \leqq p_{11} p_{22}.$$

Finally, it should be observed that the three conditions (17) are equivalent to the single condition (16).

12. Let us now assume that the two random processes considered in the preceding paragraph are *real*, i.e. that they are defined by two *real* random variables $X_1(t)$ and $X_2(t)$.

Any correlation function $R_{\mu\nu}(t)$ being in this case real, its expression as a Fourier-Stieltjes integral may be written

$$\begin{aligned} R_{\mu\nu}(t) &= \int_{-\infty}^{\infty} e^{itx}\, dF_{\mu\nu}(x) \\ &= \int_0^{\infty} [\cos tx\, df_{\mu\nu}(x) + \sin tx\, dg_{\mu\nu}(x)], \end{aligned}$$

where $f_{\mu\nu}(x)$ and $g_{\mu\nu}(x)$ are *real* functions of bounded variation in $(0, \infty)$. We may, in fact, take

$$\begin{aligned} f_{\mu\nu}(x) &= \Re[F_{\mu\nu}(x) + F_{\mu\nu}(-x)], \\ g_{\mu\nu}(x) &= \Im[-F_{\mu\nu}(x) + F_{\mu\nu}(-x)]. \end{aligned}$$

Since F_{11} and F_{22} are real and never decreasing, we then have $g_{11} = g_{22} = 0$, while f_{11} and f_{22} are never decreasing. $f_{\mu\nu}$ and $g_{\mu\nu}$ may be resolved into components of the types (a), (d) and (s) in the same way as $F_{\mu\nu}$ according to (15). Theorem 3 now immediately yields the following result.

THEOREM 4. *Four given functions $R_{\mu\nu}(t)$, $(\mu, \nu = 1, 2)$, are the correlation functions of two real continuous stationary processes if, and only if, the $R_{\mu\nu}(t)$ are given by the expressions*

$$R_{11}(t) = \int_0^\infty \cos tx \, df_{11}(x),$$

$$R_{22}(t) = \int_0^\infty \cos tx \, df_{22}(x),$$

$$R_{12}(t) = R_{21}(-t) = \int_0^\infty [\cos tx \, df_{12}(x) + \sin tx \, dg_{12}(x)],$$

where the $f_{\mu\nu}$ and $g_{\mu\nu}$ are real functions of bounded variation in $(0, \infty)$ such that f_{11} and f_{22} are never decreasing, while

$$(\Delta f_{12}^{(a)})^2 + (\Delta g_{12}^{(a)})^2 \leqq \Delta f_{11}^{(a)} \cdot \Delta f_{22}^{(a)}, \tag{18a}$$

$$(\Delta f_{12}^{(d)})^2 + (\Delta g_{12}^{(d)})^2 \leqq \Delta f_{11}^{(d)} \cdot \Delta f_{22}^{(d)}, \tag{18d}$$

$$(\Delta f_{12}^{(s)})^2 + (\Delta g_{12}^{(s)})^2 \leqq \Delta f_{11}^{(s)} \cdot \Delta f_{22}^{(s)}, \tag{18s}$$

for any interval (α, β) with $0 < \alpha < \beta$.

As in the preceding paragraph, the conditions (18a) and (18d) are equivalent to the conditions

$$(f_{12}')^2 + (g_{12}')^2 \leqq f_{11}' f_{22}'$$

and

$$p_{12}^2 + q_{12}^2 \leqq p_{11} p_{22}$$

respectively, $p_{\mu\nu}$ and $q_{\mu\nu}$ denoting here the saltuses of $f_{\mu\nu}$ and $g_{\mu\nu}$ at any given point x.

Finally, the three conditions (18s) are evidently equivalent to the single condition

$$(\Delta f_{12})^2 + (\Delta g_{12})^2 \leqq \Delta f_{11} \cdot \Delta f_{22}.$$

V. On Discrete Stationary Processes

13. Let us consider a sequence of (real or complex) random variables $X(t)$, where $t = 0, \pm 1, \pm 2, \cdots$, such that the simultaneous probability distributions of the variables $X(t_1), \cdots, X(t_k)$ are determined for any finite set of integral $t_1, \cdots, t_k$. If these probability distributions satisfy the conditions of symmetry and consistency stated in paragraph 1, we shall say that the sequence $X(t)$ defines a *discrete random process.*

Let, now, the variables $X_1(t), \cdots, X_n(t)$ define n discrete random processes such that for all integral values of t and u

$$E\{X_\mu(t)\} = 0, \qquad E\{X_\mu(t)\overline{X_\nu(u)}\} = R_{\mu\nu}(t-u), \tag{19}$$

so that the latter mean value is a function of the difference $t - u$. We shall then say that the n processes defined by the $X_\mu(t)$ are *stationary* and *stationarily correlated*. For the general theory of the discrete stationary process, reference may be made to Wold [8].

In this case, our correlation functions $R_{\mu\nu}(t)$ become sequences of numbers, depending on the (positive or negative) integral-valued variable t. Without restriction of generality we may assume

$$R_{\mu\mu}(0) = \sigma_\mu^2 = 1. \tag{20}$$

The quantity $R_{\mu\nu}(t)$ is then, in the particular case when $X_\mu(t)$ and $X_\nu(t)$ are real, identical with the ordinary correlation coefficient between $X_\mu(t_0 + t)$ and $X_\nu(t_0)$. Even in the case of complex variables $X_\mu(t)$, we shall call the $R_{\mu\nu}(t)$ the correlation coefficients of our n processes. For $\mu = \nu$, $R_{\mu\mu}(t)$ yields the *autocorrelation coefficients* of the process defined by $X_\mu(t)$, while $R_{\mu\nu}(t)$ and $R_{\nu\mu}(t)$ constitute, for $\mu \neq \nu$, the *mutual correlation coefficients* of the processes connected with $X_\mu(t)$ and $X_\nu(t)$. We obviously still have $R_{\mu\nu}(t) = \overline{R_{\nu\mu}(-t)}$.

For the correlation coefficients of a group of n discrete stationary processes, we can prove theorems which are completely analogous to our above theorems for the continuous case. In fact, the only modification required is to replace in the Fourier-Stieltjes integral (5) the infinite interval of integration by an interval of length 2π and, at the same time, to restrict the variable t to integral values. Theorem 1 will then be replaced by the following theorem.

THEOREM 5. *a) Let $X_1(t), \cdots, X_n(t)$, where $t = 0, \pm 1, \pm 2, \cdots$, be n sequences of complex random variables attached to n discrete stationary processes satisfying the conditions* (19) *and* (20). *Then the correlation coefficients $R_{\mu\nu}(t)$ defined by* (19) *are for all positive and negative integral t given by the Fourier-Stieltjes integrals*

$$R_{\mu\nu}(t) = \int_{-\pi}^{\pi} e^{itz}\, dF_{\mu\nu}(x), \tag{21}$$

where the $F_{\mu\nu}(x)$ are functions of bounded variation in $(-\pi, \pi)$, which we may always assume to be everywhere continuous to the right. Further, the form $H(z_1, \cdots, z_n)$ defined by (6) *is, for any closed interval (α, β) with $-\pi \leqq \alpha < \beta \leqq \pi$, a non-negative Hermite form. (In particular, it follows that $F_{\mu\mu}(x)$ is real and never decreasing, with a total variation equal to unity.)*

b) Conversely, let $R_{\mu\nu}(t)$, for $t = 0, \pm 1, \pm 2, \cdots$, be a set of n^2 sequences of quantities given by the Fourier-Stieltjes integrals (21), *the functions $F_{\mu\nu}(x)$ being of bounded variation in $(-\pi, \pi)$ and everywhere continuous to the right, and suppose that $R_{\mu\mu}(0) = 1$ and that the form $H(z_1, \cdots, z_n)$ is, for any closed interval*

(α, β) with $-\pi \leqq \alpha < \beta \leqq \pi$, a non-negative Hermite form. Then it is possible to find n discrete stationary processes having the correlation coefficients $R_{\mu\nu}(t)$.

The proof of this theorem is quite similar to the proof of Theorem 1. We only have to replace the definition (7) of the auxiliary function $g(x, A)$ by

$$g_{\mu\nu}(x, A) = \frac{1}{2\pi A} \sum_{t=0}^{A} \sum_{u=0}^{A} R_{\mu\nu}(t-u) e^{-ix(t-u)}.$$

According to Hausdorff [5], a necessary and sufficient condition for a representation of $R_{\mu\nu}(t)$ of the form (21), where $F_{\mu\nu}(x)$ is of bounded variation in $(-\pi, \pi)$, is

$$\int_{-\pi}^{\pi} |g_{\mu\nu}(x, A)| \, dx < \text{const.}$$

for all $A > 1$. Moreover, a necessary and sufficient condition for a representation of the form (21) with a real and never decreasing $F_{\mu\nu}(x)$ is that $g_{\mu\nu}(x, A)$ should be real and non-negative for all real x and all $A > 1$. These conditions are, of course, completely analogous to those given above for the case of a continuous variable t and a representation of the form (5), and all the rest of the proof runs on quite similar lines.

In order to obtain, for the case of discrete processes, analogues of our theorems 2, 3 and 4, we only have to replace everywhere the infinite Fourier-Stieltjes integral (5) by the corresponding finite integral (21). In theorems 3 and 4, the interval $(0, \infty)$ is then replaced by $(0, \pi)$. Otherwise everything remains unchanged.

UNIVERSITY OF STOCKHOLM,
STOCKHOLM, SWEDEN

REFERENCES

[1] H. Cramér, *Random variables and probability distributions*, Cambridge Tracts in Mathematics, No. 36, Cambridge 1937.

[2] ———, *On the representation of functions by certain Fourier integrals*, Trans. Amer. Math. Soc., vol. 46 (1939), p. 191.

[3] J. L. Doob, *Stochastic processes depending on a continuous parameter*, Trans. Amer. Math. Soc., vol. 42 (1937), p. 107.

[4] M. Fréchet, *Recherches théoriques modernes sur la théorie des probabilités*, premier livre, Paris 1937.

[5] F. Hausdorff, *Momentprobleme für ein endliches Intervall*, Math. Zeitschrift, vol. 16 (1923), p. 220.

[6] A. Khintchine, *Korrelationstheorie der stationären stochastischen Prozesse*, Math. Annalen, vol. 109 (1934), p. 604.

[7] P. Lévy, *Théorie de l'addition des variables aléatoires*, Paris, 1937.

[8] H. Wold, *A study in the analysis of stationary time series*, Inaugural Dissertation (Stockholm), Uppsala 1938.

ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

COHOMOLOGY AND CONTINUOUS MAPPINGS

BY SAMUEL EILENBERG

(Received April 25, 1939)

I. INTRODUCTION[1]

Given a geometrical cell-complex K and a polyhedron Y, let us consider all continuous mappings $f(K) \subset Y$. These maps are divided into homotopy classes, the maps in any class being homotopic to each other. The problem of determining these classes by means of known invariants (for instance, homology properties) of K and Y is of extremely great importance in modern topology.

The discussion of the case when $K = S^i$ is an i-dimensional spherical manifold led Hurewicz[2] to the definition of the i^{th} homotopy group $\pi_i(Y)$ of Y. Although our knowledge of these groups is still very imperfect, they have proved to be a powerful instrument in all considerations connected with the general question.

Special interest has always been paid by topologists to the case when $Y = S^n$. In this direction there is the fundamental result of Hopf solving the question completely if $K = K^n$.[3] As newly shown by Whitney[4] this theorem may be stated as follows:

The classes of maps of K^n into S^n are in a $(1 - 1)$-correspondence with the elements of the n^{th} cohomology group ${}^n\boldsymbol{H}_I(K^n)$ of K^n with the group I of all integers as coefficient group.

The theorem holds even if we replace the condition $Y = S^n$ by the condition $\pi_i(Y) = 0$ for $i < n$ $(n > 1)$, provided the group I is replaced by $\pi_n(Y)$. This shows that the appearance of the group I in Hopf's theorem (as stated above) is due to the fact that I and $\pi_n(S^n)$ are isomorphic.

The use of cohomology in Hopf's theorem is natural, also, for the following reason. The theorem and the proofs hold for infinite, locally finite complexes as well as for finite, provided infinite chains and cocycles are admitted.[5] The statement of Hopf's theorem for infinite complexes in the language of homology seems to be much more complicated.[6]

---

[1] Some of the results were published by the author without proofs, in C. R. Paris 208 (1939), p. 68. See also H. Freudenthal, Proc. Akad. Amsterdam 42(1939), p. 139.

[2] W. Hurewicz, Proc. Akad. Amsterdam 38(1935), pp. 112–115.

[3] K^n stands for the closed subcomplex of K consisting of all its cells of dimension $\leqq n$; $K^{-1} = 0$.

[4] H. Whitney, Duke Math. Jour. 3(1937), p. 51. References to Hopf, Hurewicz, and others will be found there.

[5] C. H. Dowker, Proc. Nat. Acad. U. S. A. 23(1937), p. 293.

[6] In a note published in C. R. Paris 206(1938), p. 1436, L. Pontrjagin signalizes some results concerning the case $K = K^{n+1}$, $Y = S^n$. $(n + 1)$-chains with coefficients from $\pi_{n+1}(S^n)$ are implicitly introduced.

In a recent paper[7] I have made systematic use of chains with elements of homotopy groups as coefficients. Let Q^r be a geometrical cell-complex which is an r-dimensional oriented combinatorial manifold, and let P^i be an i-dimensional closed subcomplex of Q^r. Given a map $f(Q^r - P^i) \subset S^n$, I define an i-chain $\gamma^i(f)$ in P^i with coefficients in $\pi_{r-i-1}(S^n)$ as follows: Let σ^i be an i-cell of P^i and s^{r-i-1} a "small" $(r - i - 1)$-sphere contained in $Q^r - P^i$ and "simply linked" with σ^i, then the element of $\pi_{r-i-1}(S^n)$ defined by the map $f(s^{r-i-1}) \subset S^n$ is the coefficient of σ^i in $\gamma^i(f)$. I proved that this chain is a cycle and that its homology properties are closely related with the homotopy properties of f.

The purpose of the present paper is to build up an analogous theory for arbitrary complexes. In parts III, IV, and V we develop such a theory which, applied to manifolds, will at once give us, in part VI, all previous results concerning $\gamma^i(f)$, even in a slightly stronger form.

The basic definition is introduced in part III as follows: Let K be an arbitrary geometrical cell-complex and $f(K^n) \subset Y$ a continuous mapping. Let σ^{n+1} be a $(n + 1)$-cell of K and $c(f, \sigma^{n+1})$ the element of $\pi_n(Y)$ defined by considering f on the boundary of σ^{n+1}. Taking $c(f, \sigma^{n+1})$ as the coefficient of σ^{n+1}, we obtain an $(n + 1)$-chain $c^{n+1}(f)$. It is proved that $c^{n+1}(f)$ is a cocycle. Extension theorem I, which is the main theorem of this paper, shows how closely the cohomology properties of this cocycle are connected with the extension-possibilities of f. In all of part III Y can be an arbitrary topological space which is simple in dimension n, a condition introduced in part II that is necessary in order to make the definition of $c(f, \sigma^{n+1})$ unique.

Part IV contains the application to the case when $\pi_i(Y) = 0$ for $i < n$. A generalization of Hopf's theorem is given which includes the generalization of Hurewicz-Whitney[4] and a generalization given by the author[8] arising from replacing the hypothesis $K = K^n$ by some hypothesis concerning cohomology groups for dimensions $> n$.

A homology interpretation of the results of part IV is given in part V, under some additional hypothesis on Y.

Two appendices discussing special topics are given at the end of the paper.

II. Preliminaries

1. Let K be a geometrical locally finite complex, with oriented convex cells σ_i^n of dimension $n = 0, 1, \cdots$. Their number may be infinite and their dimensions may form an unbounded sequence. The cells are open, and the closure of the n-cell σ_i^n will be denoted by $\bar{\sigma}_i^n$.

Let $\partial_{ji}^n = 1, -1$ or 0 according as σ_j^{n-1} is positively, negatively, or not at all, on the boundary $\bar{\sigma}_i^n - \sigma_i^n$ of σ_i^n. The *boundary* and *coboundary* of σ_i^n are defined by

$$\partial\sigma_i^n = \sum_j \partial_{ji}^n \sigma_j^{n-1}, \qquad \delta\sigma_i^n = \sum_j \partial_{ij}^{n+1} \sigma_j^{n+1}.$$

[7] Fund. Math. 31(1938), pp. 179–200.

[8] Compositio Math. 6(1939), p. 429.

An n-chain is an infinite linear form $A^n = \sum_i \alpha_i \sigma_i^n$, in which α_i are elements of an abelian group G[9]. The *boundary* and *coboundary* of A^n are defined by

$$\partial A^n = \sum_i \alpha_i \partial \sigma_i^n, \qquad \delta A^n = \sum_i \alpha_i \delta \sigma_i^n.$$

2. Let K' be a closed subcomplex of K. An n-chain $A^n = \sum_i \alpha_i \sigma_i^n$ is contained in K' or in $K - K'$ (notation: $A^n \subset K'$ or $A^n \subset K - K'$) if $\alpha_i = 0$ for each n-cell σ_i^n in $K - K'$ or in K'. Clearly $A^n \subset K'$ implies $\partial A^n \subset K'$ and $A^n \subset K - K'$ implies $\delta A^n \subset K - K'$.

A^n is a *cycle* mod K' if $\partial A^n \subset K'$; A^n is a *cocycle in* $K - K'$ if $A^n \subset K - K'$ and $\delta A^n = 0$. Two cycles A_0^n and A_1^n mod K' are *homologous* mod K' (notation: $A_0^n \sim A_1^n \bmod K'$) if there is an $(n+1)$-chain A^{n+1} such that $\partial A^{n+1} - A_1^n - A_0^n \subset K$. Two cocycles A_0^n and A_1^n in $K - K'$ are *cohomologous in* $K - K'$ (notation: $A_0^n \backsim A_1^n$ in $K - K'$) if there is an $(n-1)$-chain $A^{n-1} \subset K - K'$ such that $\delta A^{n-1} = A_0^n - A_1^n$.

Using the relations $\partial\partial A^n = 0 = \delta\delta A^n$ we may define as usual the *homology* and *cohomology groups*

$$^n\boldsymbol{H}^G(K) \bmod K', \qquad ^n\boldsymbol{H}_G(K - K'),$$

where G is an arbitrary abelian group whose elements are taken as coefficients in the chains.

If $K' = 0$ we write $A_0^n \sim A_1^n$, $A_0^n \backsim A_1^n$ and $^n\boldsymbol{H}^G(K)$ instead of $A_0^n \sim A_1^n \bmod K'$, $A_0^n \backsim A_1^n$ in $K - K'$ and $^n\boldsymbol{H}^G(K) \bmod K'$.

Everything can be repeated starting from *finite* chains. We shall use similar notations, replacing $\sim$, $\backsim$ and $\boldsymbol{H}$ by $\sim^*$, $\backsim^*$ and $\boldsymbol{H}^*$.

3. LEMMA. *Given an n-chain A^n, an $(n+1)$-cell σ^{n+1} and an n-cell $\sigma^n \subset \bar{\sigma}^{n+1}$, there is an $(n-1)$-chain A^{n-1} such that*

$$A^n - \delta A^{n-1} = \alpha\sigma^n + A_1^n, \qquad A_1^n \subset K - \bar{\sigma}^{n+1}, \qquad \alpha \in G. \tag{3.1}$$

PROOF. Let $\sigma^n, \sigma_1^n, \sigma_2^n, \cdots, \sigma_r^n$ be all the n-cells of $\bar{\sigma}^{n+1}$ and let

$$A^n = \beta\sigma^n + \sum_{i=1}^{r} \beta_i \sigma_i^n + B^n, \qquad B^n \subset K - \bar{\sigma}^{n+1}.$$

The boundary $\bar{\sigma}^{n+1} - \sigma^{n+1}$ of σ^{n+1} being an n-dimensional manifold there is[10] an $(n-1)$-chain A_i^{n-1} with integer coefficients such that

$$\delta A_i^{n-1} = \sigma_i^n - \epsilon_i \sigma^n + B_i^n, \qquad \epsilon_i = \pm 1, \qquad B_i^n \subset K - \bar{\sigma}^{n+1}$$

Writing $A^{n-1} = \sum_{i=1}^{r} \beta_i A_i^{n-1}$, $\alpha = \beta + \sum_{i=1}^{r} \epsilon_i \beta_i$ and $A_1^n = B^n + \sum_{i=1}^{r} \beta_i B_i^n$ we obtain (3.1).

---

[9] More exactly, an n-chain is a function with n-cells as arguments and elements of G as values.

[10] H. Whitney, Duke Math. Jour. 3(1937), p. 44.

In a recent paper[7] I have made systematic use of chains with elements of homotopy groups as coefficients. Let Q^r be a geometrical cell-complex which is an r-dimensional oriented combinatorial manifold, and let P^i be an i-dimensional closed subcomplex of Q^r. Given a map $f(Q^r - P^i) \subset S^n$, I define an i-chain $\gamma^i(f)$ in P^i with coefficients in $\pi_{r-i-1}(S^n)$ as follows: Let σ^i be an i-cell of P^i and s^{r-i-1} a "small" $(r - i - 1)$-sphere contained in $Q^r - P^i$ and "simply linked" with σ^i, then the element of $\pi_{r-i-1}(S^n)$ defined by the map $f(s^{r-i-1}) \subset S^n$ is the coefficient of σ^i in $\gamma^i(f)$. I proved that this chain is a cycle and that its homology properties are closely related with the homotopy properties of f.

The purpose of the present paper is to build up an analogous theory for arbitrary complexes. In parts III, IV, and V we develop such a theory which, applied to manifolds, will at once give us, in part VI, all previous results concerning $\gamma^i(f)$, even in a slightly stronger form.

The basic definition is introduced in part III as follows: Let K be an arbitrary geometrical cell-complex and $f(K^n) \subset Y$ a continuous mapping. Let σ^{n+1} be a $(n + 1)$-cell of K and $c(f, \sigma^{n+1})$ the element of $\pi_n(Y)$ defined by considering f on the boundary of σ^{n+1}. Taking $c(f, \sigma^{n+1})$ as the coefficient of σ^{n+1}, we obtain an $(n + 1)$-chain $c^{n+1}(f)$. It is proved that $c^{n+1}(f)$ is a cocycle. Extension theorem I, which is the main theorem of this paper, shows how closely the cohomology properties of this cocycle are connected with the extension-possibilities of f. In all of part III Y can be an arbitrary topological space which is simple in dimension n, a condition introduced in part II that is necessary in order to make the definition of $c(f, \sigma^{n+1})$ unique.

Part IV contains the application to the case when $\pi_i(Y) = 0$ for $i < n$. A generalization of Hopf's theorem is given which includes the generalization of Hurewicz-Whitney[4] and a generalization given by the author[8] arising from replacing the hypothesis $K = K^n$ by some hypothesis concerning cohomology groups for dimensions $> n$.

A homology interpretation of the results of part IV is given in part V, under some additional hypothesis on Y.

Two appendices discussing special topics are given at the end of the paper.

II. Preliminaries

1. Let K be a geometrical locally finite complex, with oriented convex cells σ_i^n of dimension $n = 0, 1, \cdots$. Their number may be infinite and their dimensions may form an unbounded sequence. The cells are open, and the closure of the n-cell σ_i^n will be denoted by $\bar{\sigma}_i^n$.

Let $\partial_{ji}^n = 1, -1$ or 0 according as σ_j^{n-1} is positively, negatively, or not at all, on the boundary $\bar{\sigma}_i^n - \sigma_i^n$ of σ_i^n. The *boundary* and *coboundary* of σ_i^n are defined by

$$\partial\sigma_i^n = \sum_j \partial_{ji}^n \sigma_j^{n-1}, \qquad \delta\sigma_i^n = \sum_j \partial_{ij}^{n+1} \sigma_j^{n+1}.$$

[7] Fund. Math. 31(1938), pp. 179–200.

[8] Compositio Math. 6(1939), p. 429.

An n-chain is an infinite linear form $A^n = \sum_i \alpha_i \sigma_i^n$, in which α_i are elements of an abelian group G[9]. The *boundary* and *coboundary* of A^n are defined by

$$\partial A^n = \sum_i \alpha_i \partial \sigma_i^n, \qquad \delta A^n = \sum_i \alpha_i \delta \sigma_i^n.$$

2. Let K' be a closed subcomplex of K. An n-chain $A^n = \sum_i \alpha_i \sigma_i^n$ is contained in K' or in $K - K'$ (notation: $A^n \subset K'$ or $A^n \subset K - K'$) if $\alpha_i = 0$ for each n-cell σ_i^n in $K - K'$ or in K'. Clearly $A^n \subset K'$ implies $\partial A^n \subset K'$ and $A^n \subset K - K'$ implies $\delta A^n \subset K - K'$.

A^n is a *cycle* mod K' if $\partial A^n \subset K'$; A^n is a *cocycle in* $K - K'$ if $A^n \subset K - K'$ and $\delta A^n = 0$. Two cycles A_0^n and A_1^n mod K' are *homologous* mod K' (notation: $A_0^n \sim A_1^n \bmod K'$) if there is an $(n+1)$-chain A^{n+1} such that $\partial A^{n+1} - A_1^n - A_0^n \subset K$. Two cocycles A_0^n and A_1^n in $K - K'$ are *cohomologous in* $K - K'$ (notation: $A_0^n \backsim A_1^n$ in $K - K'$) if there is an $(n-1)$-chain $A^{n-1} \subset K - K'$ such that $\delta A^{n-1} = A_0^n - A_1^n$.

Using the relations $\partial\partial A^n = 0 = \delta\delta A^n$ we may define as usual the *homology* and *cohomology groups*

$$^nH^G(K) \bmod K', \qquad ^nH_G(K - K'),$$

where G is an arbitrary abelian group whose elements are taken as coefficients in the chains.

If $K' = 0$ we write $A_0^n \sim A_1^n$, $A_0^n \backsim A_1^n$ and $^nH^G(K)$ instead of $A_0^n \sim A_1^n \bmod K'$, $A_0^n \backsim A_1^n$ in $K - K'$ and $^nH^G(K) \bmod K'$.

Everything can be repeated starting from *finite* chains. We shall use similar notations, replacing $\sim$, $\backsim$ and H by $\sim^*$, $\backsim^*$ and H^*.

3. LEMMA. *Given an n-chain A^n, an $(n+1)$-cell σ^{n+1} and an n-cell $\sigma^n \subset \bar{\sigma}^{n+1}$, there is an $(n-1)$-chain A^{n-1} such that*

$$A^n - \delta A^{n-1} = \alpha\sigma^n + A_1^n, \qquad A_1^n \subset K - \bar{\sigma}^{n+1}, \qquad \alpha \in G. \tag{3.1}$$

PROOF. Let $\sigma^n, \sigma_1^n, \sigma_2^n, \cdots, \sigma_r^n$ be all the n-cells of $\bar{\sigma}^{n+1}$ and let

$$A^n = \beta\sigma^n + \sum_{i=1}^{r} \beta_i \sigma_i^n + B^n, \qquad B^n \subset K - \bar{\sigma}^{n+1}.$$

The boundary $\bar{\sigma}^{n+1} - \sigma^{n+1}$ of σ^{n+1} being an n-dimensional manifold there is[10] an $(n-1)$-chain A_i^{n-1} with integer coefficients such that

$$\delta A_i^{n-1} = \sigma_i^n - \epsilon_i \sigma^n + B_i^n, \qquad \epsilon_i = \pm 1, \qquad B_i^n \subset K - \bar{\sigma}^{n+1}$$

Writing $A^{n-1} = \sum_{i=1}^{r} \beta_i A_i^{n-1}$, $\alpha = \beta + \sum_{i=1}^{r} \epsilon_i \beta_i$ and $A_1^n = B^n + \sum_{i=1}^{r} \beta_i B_i^n$ we obtain (3.1).

---

[9] More exactly, an n-chain is a function with n-cells as arguments and elements of G as values.

[10] H. Whitney, Duke Math. Jour. 3(1937), p. 44.

4. Let X and Y be two topological spaces and Y^X the family of all continuous transformations f such that $f(X) \subset Y$. Given $f \in Y^X$ and $A \subset X$, we denote by $f \mid A$ the "partial" function obtained considering f only on A. $Y^X(A, f)$ will be the subfamily of Y^X containing all the functions f' such that $f' \mid A = f \mid A$.

Let us fix a point $y_0 \in Y$. We shall denote by 0 the function mapping the whole of X into y_0, 0 will also stand for $0 \mid A$ and $Y^X(A, y_0)$ for $Y^X(A, 0)$.

We shall denote the closed interval (0, 1) by E and the Cartesian product of X and E by $X \times E$.

Two functions $f_0, f_1 \in Y^X(A, f)$ will be called *homotopic relative to* A (notation: $f_0 \simeq f_1$ rel. A) if there is a $g \in Y^{X \times E}$ such that

$$g(x, i) = f_i(x) \quad \text{for} \quad x \in X, i = 0, 1,$$
$$g(x, t) = f(x) \quad \text{for} \quad x \in A, t \in E.$$

If $A = 0$, the functions f_0 and f_1 will merely be called *homotopic* (notation: $f_0 \simeq f_1$) instead of homotopic rel. 0.

In this way the family Y^X is divided into *homotopy classes* and $Y^X(A, f)$ into *homotopy classes rel.* A.

5. It is well known[11] that the set $T = K \times 0 + K' \times E$ is a retract[12] of the product $K \times E$, where K' is a closed subcomplex of K, and therefore

(5.1) *Every* $f \in Y^T$ *has an extension* $f' \in Y^{K \times E}$.

Let $A \subset K'$ and $f \in Y^K$.

(5.2) *Given* $f_0 \in Y^K(A, f)$ *such that* $f_0 \mid K' \simeq f \mid K'$ *rel.* A, *there is an* $f_0' \in Y^K(K', f)$ *such that* $f_0' \simeq f_0$ *rel.* A.

PROOF. Let $g \in Y^{K' \times E}$ be such that

$$g(x, 0) = f_0(x), \; g(x, 1) = f(x) \text{ for } x \in K',$$
$$g(x, t) = f(x) \quad \text{for } x \in A, t \in E.$$

Writing
$$g'(x, t) = g(x, t) \text{ for } (x, t) \in K' \times E,$$
$$g'(x, t) = f_0(x) \quad \text{for } (x, t) \in K \times 0$$

we have $g' \in Y^T$ and by (5.1) there is an extension $g'' \in Y^{K \times E}$ of g'. The function $f_0'(x) = g''(x, 1)$ satisfies the conditions of (5.2).

Taking $A = 0$ in (5.2) we have

(5.3) *Given* $f_0, f_1 \in Y^K$ *such that* $f_0 \mid K' \simeq f_1 \mid K'$, *there is an* $f_0' \in Y^K$ *such that* $f_0' \simeq f_0$ *and* $f_0' \mid K' = f_1 \mid K'$.

6. Let S^n be an oriented n-dimensional sphere, let $S^n = E_+^n + E_-^n$ be a decomposition of S^n into two hemispheres (oriented as S^n), and let x_0 be a point of the equator $S^{n-1} = E_+^n \cdot E_-^n$.

---

[11] See P. Alexandroff and H. Hopf, *Topologie I*, Berlin 1935, p. 501; K. Borsuk, Ann. Soc. Polon. Math. 16(1937), p. 218.

[12] $A \subset X$ is a *retract* of X if there is an $r \in A^X$ such that $r(x) = x$ for $x \in A$.

We shall denote by $[f]$ the homotopy class rel. x_0 of a map $f \in Y^{S_n}(x_0, y_0)$. Given $f_0, f_1 \in Y^{S_n}(x_0, y_0)$, we obviously have

$$f_0 \mid E^n_- \simeq 0 \text{ rel. } x_0 \quad \text{and} \quad f_1 \mid E^n_+ \simeq 0 \text{ rel. } x_0 .$$

Therefore, by (5.2), there are two functions $f_0' \in [f_0]$ and $f_1' \in [f_1]$ such that

$$f_0' \mid E^n_- = 0 \quad \text{and} \quad f_1' \mid E^n_+ = 0.$$

Writing

$$f \mid E^n_+ = f_0' \mid E^n_+ \quad \text{and} \quad f \mid E^n_- = f_1' \mid E^n_-$$

we have $f \in Y^{S_n}(x_0, y_0)$. We define

$$[f] = [f_0] + [f_1].$$

It is easy to see that with this definition the homotopy classes rel. x_0 of $Y^{S_n}(x_0, y_0)$ form a group $\pi_n = \pi_n(Y)$ which is called[13] the n^{th} *homotopy group* of Y (with respect to y_0). The unit element of this group is obviously the class $[0]$.

(6.1) *π_n is abelian for $n > 1$.*[14]

7. We shall call Y *simple in dimension n*,[15] or, more briefly, *n-simple* if every homotopy class of Y^{S_n} contains exactly one homotopy class rel. x_0 of $Y^{S_n}(x_0, y_0)$. In other words, Y is n-simple if

(a) *for each $f \in Y^{S_n}$ there is an $f' \in Y^{S_n}(x_0, y_0)$ such that $f' \simeq f$*

(b) *$f_0, f_1 \in Y^{S_n}(x_0, y_0)$ and $f_0 \simeq f_1$ imply $f_0 \simeq f_1$ rel. x_0.*

It can easily be shown, using (5.3), that (a) is equivalent with

(a′) *Y is arcwise connected.*

From now on we shall assume that Y is arcwise connected.

Obviously if Y is n-simple then every $f \in Y^{S_n}$ determines uniquely an element of π_n, and we may suppose that the elements of π_n are the homotopy classes of Y^{S_n}. This is the property that makes n-simple spaces useful for our further discussion.

(7.1) *π_1 is abelian if and only if Y is 1-simple.*[15]

(7.2) *If $\pi_n = 0$ then Y is n-simple.*[15]

(7.3) *If $\pi_1 = 0$ then Y is n-simple.*[15]

(7.4) *S^r is n-simple for $r = 1, 2, \cdots$.*

8. Let E^n be an oriented n-dimensional element bounded by S^{n-1}. Let us choose two homeomorphisms

$$h_+(E^n_+) = E^n \quad \text{and} \quad h_-(E^n_-) = E^n$$

---

[13] W. Hurewicz, Proc. Akad. Amsterdam 38(1935), p. 113.

[14] Ibid., p. 114. Though π_1 is in general non-abelian, it is abelian in all the cases considered here, so that additive notations will be used throughout this paper.

[15] S. Eilenberg, Fund. Math. 32(1939), pp. 167–175.

such that $h_+ \mid S^{n-1} = h_- \mid S^{n-1}$, $S^{n-1} = E_+^n \cdot E_-^n$ being the equator of S^n. Let us agree further that h_+ transforms E_+^n into E^n positively. Then h_- transforms E_-^n into E^n negatively.

Given two functions $f_0, f_1 \in Y^{E^n}$ such that $f_0 \mid S^{n-1} = f_1 \mid S^{n-1}$, let $(f_0, f_1) \in Y^{S_n}$ be a function defined by

$$(f_0, f_1) \mid E_+^n = f_0 h_+, \qquad (f_0, f_1) \mid E_-^n = f_1 h_-.$$

Assuming that Y is n-simple we denote by $d(f_0, f_1)$ the element of π_n corresponding to (f_0, f_1).

(8.1) $d(f_0, f_1) = 0$ *if and only if* $f_0 \simeq f_1$ *rel.* S^{n-1}.

(8.2) *If* $f_0, f_1, f_2 \in Y^{E^n}$ *and* $f_0 \mid S^{n-1} = f_1 \mid S^{n-1} = f_2 \mid S^{n-1}$ *then* $d(f_0, f_1) + d(f_1, f_2) = d(f_0, f_2)$,

(8.3) $d(f_0, f_1) = -d(f_1, f_0)$.

(8.4) *Given* $f_0 \in Y^{E^n}$ *and* $\alpha \in \pi_n$ *there is an* $f_1 \in Y^{E^n}$ *such that* $f_0 \mid S^{n-1} = f_1 \mid S^{n-1}$ *and that* $d(f_0, f_1) = \alpha$.

Proofs. Ad (8.1). $d(f_0, f_1) = 0$ is equivalent to $(f_0, f_1) \simeq 0$ and this obviously holds if and only if $f_0 \simeq f_1$ rel. S^{n-1}.

Ad (8.2). Since $f_1 \simeq 0$ there is a map $g_1 \in Y^{E^n \times E}$ such that

$$g_1(x, 0) = f_1(x), \qquad g_1(x, 1) = y_0 \quad \text{for} \quad x \in E^n$$

using (5.1) we can find two maps $g_0, g_2 \in Y^{E^n \times E}$ such that

$$g_i(x, 0) = f_i(x) \quad \text{for} \quad x \in E^n, \qquad i = 0, 2,$$

$$g_i(x, t) = g_1(x, t) \quad \text{for} \quad (x, t) \in S^{n-1} \times E, \qquad i = 0, 2.$$

Writing $f_{i,t}(x) = g_i(x, t)$ for $t \in E$ and $i = 0, 1, 2$ we have for $i, j = 0, 1, 2$

$$f_{i,t} \mid S^{n-1} = f_{j,t} \mid S^{n-1} \quad \text{and} \quad (f_i, f_j) \simeq (f_{i,t}, f_{j,t}).$$

In particular we have $d(f_i, f_j) = d(f_{i,1}, f_{j,1})$ and since $f_{1,1} = 0$, (8.2) reduces to the formula

$$d(f_{0,1}, 0) + d(0, f_{2,1}) = d(f_{0,1}, f_{2,1}),$$

which is a direct consequence of the definition of π_n.

Ad (8.3). Follows from (8.1) and (8.2) taking $f_2 = f_0$.

Ad (8.4). Let $g \in Y^{S_n}$ be a map representing the element $\alpha \in \pi_n$. Since $g \mid E_+^n \simeq 0$ and $f_0 \simeq 0$ we have

$$g \mid E_+^n \simeq f_0 h_+.$$

Using (5.3) we find a $g' \in Y^{S^n}$ such that

$$g' \simeq g \quad \text{and} \quad g' \mid E_+^n = f_0 h_+.$$

Writing $f_1 = g' h_-^{-1}$ we have $f_1 \in Y^{E^n}$, $f_0 \mid S^{n-1} = f_1 \mid S^{n-1}$ and $g' = (f_0, f_1)$. Therefore $d(f_0, f_1) = \alpha$.

III. The General Theory

9. We assume that Y is n-simple. All chains considered here will have coefficients taken from $\pi_n = \pi_n(Y)$.

Let $f \in Y^{K^n}$.[3] For each $(n+1)$-cell σ^{n+1} the map $f \mid (\sigma^{n+1})^\cdot$, the n-sphere $(\sigma^{n+1})^\cdot = \bar{\sigma}^{n+1} - \sigma^{n+1}$ being oriented by the n-cycle $\partial\sigma^{n+1}$, defines uniquely an element of π_n which we denote by $c(f, \sigma^{n+1})$. We define an $(n+1)$-chain $c^{n+1}(f)$ writing

$$c^{n+1}(f) = \sum_i c(f, \sigma_i^{n+1})\sigma_i^{n+1}.^{16} \tag{9.1}$$

Obviously $c(f, \sigma^{n+1}) = 0$ is equivalent with $f \mid (\sigma^{n+1})^\cdot \simeq 0$ and therefore

(9.2) *$c(f, \sigma^{n+1}) = 0$ if and only if there is an extension $f' \in Y^{K^n + \sigma^{n+1}}$ of f.*

(9.3) *If $f' \in Y^{K^n}$ and $f \simeq f'$ then $c^{n+1}(f) = c^{n+1}(f')$.*

Let $f_0, f_1 \in Y^{K^n}$, $f_0 \mid K^{n-1} = f_1 \mid K^{n-1}$. For each n-cell σ^n the maps $f_0 \mid \bar{\sigma}^n$ and $f_1 \mid \bar{\sigma}^n$ define according to **8** an element $d(f_0, f_1, \sigma^n)$ of π_n. We define[17] an n-chain $d^n(f_0, f_1)$ writing

$$d^n(f_0, f_1) = \sum_i d(f_0, f_1, \sigma_i^n)\sigma_i^n. \tag{9.4}$$

It follows from (8.1) that

(9.5) *$d(f_0, f_1, \sigma^n) = 0$ if and only if $f_0 \mid \bar{\sigma}^n \simeq f_1 \mid \bar{\sigma}^n$ rel. $\bar{\sigma}^n - \sigma^n$.*

10. We shall prove the following fundamental properties of $c^{n+1}(f)$ and $d^n(f_0, f_1)$:

(10.1) *$c^{n+1}(f)$ is a cocycle* (i.e. $\delta c^{n+1}(f) = 0$).

(10.2) $\delta d^n(f_0, f_1) = c^{n+1}(f_0) - c^{n+1}(f_1)$.

(10.3) *If $f_0, f_1, f_2 \in Y^{K^n}$ and $f_0 \mid K^{n-1} = f_1 \mid K^{n-1} = f_2 \mid K^{n-1}$ then* $d^n(f_0, f_1) + d^n(f_1, f_2) = d^n(f_0, f_2)$.

(10.4) $d^n(f_0, f_1) = -d^n(f_1, f_0)$

(10.5) *Given $f_0 \in Y^{K^n}$ and an n-chain d^n there is an $f_1 \in Y^{K^n}$ such that $f_0 \mid K^{n-1} = f_1 \mid K^{n-1}$ and that $d^n(f_0, f_1) = d^n$.*

Proofs. (10.3) follows from (8.2); (10.4) follows from (8.3) and (10.5) is an immediate consequence of (8.4).

Ad (10.2). Let σ_i^n be an n-cell. We suppose first that

$$f_0 \mid K^n - \sigma_i^n = f_1 \mid K^n - \sigma_i^n. \tag{10.6}$$

We then have

$$d^n(f_0, f_1) = d(f_0, f_1, \sigma_i^n)\sigma_i^n,$$

and therefore

$$\delta d^n(f_0, f_1) = d(f_0, f_1, \sigma_i^n)\delta\sigma_i^n = d(f_0, f_1, \sigma_i^n)\sum_j \partial_{ij}^{n+1}\sigma_j^{n+1}.$$

[16] According to footnote[9] we obtain $c^{n+1}(f)$ by considering $c(f, \sigma_i^{n+1})$ as a function of the argument σ_i^{n+1}.

[17] See the paper of L. Pontrjagin quoted in [6].

The formula (10.2) therefore, takes the form

$$\partial_{ij}^{n+1} d(f_0, f_1, \sigma_i^n) = c(f_0, \sigma_j^{n+1}) - c(f_1, \sigma_j^{n+1}). \tag{10.7}$$

If $\partial_{ij}^{n+1} = 0$, then because of (10.6) we have $f_0 \mid (\sigma_j^{n+1})^{\cdot} = f_1 \mid (\sigma_j^{n+1})^{\cdot}$. It follows that $c(f_0, \sigma_j^{n+1}) = c(f_1, \sigma_j^{n+1})$, and (10.7) holds.

If $\partial_{ij}^{n+1} \neq 0$ we have $\sigma_i^n \subset \bar{\sigma}_j^{n+1}$. Let $h(\bar{\sigma}_i^n) = (\sigma_j^{n+1})^{\cdot} - \sigma_i^n$ be a homeomorphism such $h(x) = x$ for $x \in (\sigma_i^n)^{\cdot}$. Because of (10.6) we have

$$f_0 h = f_1 h. \tag{10.8}$$

It follows from **8** that for $k = 0, 1$ we have

$$c(f_k, \sigma_j^{n+1}) = d(f_k, f_k h, \sigma_i^n) \quad \text{if} \quad \partial_{ij}^{n+1} = 1,$$

$$c(f_k, \sigma_j^{n+1}) = d(f_k h, f_k, \sigma_i^n) \quad \text{if} \quad \partial_{ij}^{n+1} = -1.$$

Using (10.8) and (8.3) we obtain

$$c(f_0, \sigma_j^{n+1}) - c(f_1, \sigma_j^{n+1}) = \partial_{ij}^{n+1} [d(f_0, f_0 h, \sigma_i^n) + d(f_0 h, f_1, \sigma_i^n)].$$

(10.7) follows, therefore, from (8.2).

Now, let $\sigma_1^n, \sigma_2^n, \cdots, \sigma_r^n$ be the n-cells of $\bar{\sigma}^{n+1}$, where σ_j^{n+1} is an arbitrary $(n+1)$-cell. We define the functions $g_0, g_1, \cdots, g_r \in Y^{K^n}$ as follows:

$$g_0 = f_0,$$

$$g_i \mid K^n - \sigma_i^n = g_{i-1} \mid K^n - \sigma_i^n, \; g_i \mid \bar{\sigma}_i^n = f_1 \mid \bar{\sigma}_i^n \text{ for } i > 0.$$

We then have for $i = 0, 1, \cdots, r-1$

$$\delta d^n(g_i, g_{i+1}) = c^{n+1}(g_i) - c^{n+1}(g_{i+1}),$$

and by (10.3)

$$\delta d^n(f_0, g_r) = c^{n+1}(f_0) - c^{n+1}(g_r). \tag{10.9}$$

But since $g_r \mid (\sigma_j^{n+1})^{\cdot} = f_2 \mid (\sigma_j^{n+1})^{\cdot}$ we have

$$c(g_r, \sigma_j^{n+1}) = c(f_1, \sigma_j^{n+1}),$$

$$d(f_0, g_r, \sigma_i^n) = d(f_0, f_1, \sigma_i^n) \quad \text{for} \quad \sigma_i^n \subset \bar{\sigma}_j^{n+1}.$$

Consequently we deduce from (10.9) that $c(f_0, \sigma_j^{n+1}) - c(f_1, \sigma_j^{n+1})$ is the coefficient of σ_j^{n+1} in the $(n+1)$-chain $\delta d^n(f_0, f_1)$, and hence (10.2) is completely proved.

Ad (10.1). Let σ^{n+2} be an $(n+2)$-cell and $\sigma^{n+1} \subset \bar{\sigma}^{n+2}$ an $(n+1)$-cell. By the lemma of **3** there is an n-chain d^n such that

$$c^{n+1}(f) - \delta d^n = \alpha \sigma^{n+1} + A^{n+1}, \qquad A^{n+1} \subset K - \bar{\sigma}^{n+2}, \qquad \alpha \in \pi_n. \tag{10.10}$$

According to (10.5) there is an $f' \in Y^{K^n}$ such that $f \mid K^{n-1} = f' \mid K^{n-1}$ and $d^n(f, f') = d^n$. Therefore, by (10.2),

$$c^{n+1}(f) - \delta d^n = c^{n+1}(f'). \tag{10.11}$$

It follows from (10.10) and (10.11) that $c(f', \sigma^{n+1}) = \alpha$ and that $c(f', \sigma_i^{n+1}) = 0$ for each $(n+1)$-cell $\sigma_i^{n+1} \subset \bar{\sigma}^{n+2} - \sigma^{n+1}$. By (9.2) we may therefore suppose that f' is extended on $(\sigma^{n+2})^{\cdot} - \sigma^{n+1}$. This complex being an $(n+1)$-element it follows $f' \mid (\sigma^{n+1})^{\cdot} \simeq 0$ therefore $c(f', \sigma^{n+1}) = 0$ and therefore $\alpha = 0$. This and (10.10) imply

$$c^{n+1}(f) - \delta d^n \subset K - \bar{\sigma}^{n+2},$$

and therefore

$$\delta c^{n+1}(f) = \delta[c^{n+1}(f) - \delta d^n] \subset K - \bar{\sigma}^{n+2}.$$

σ^{n+2} being an arbitrary $(n+2)$-cell it follows that $\delta c^{n+1}(f) = 0$.

11. Let $f \in Y^{K'+K^n}$, where K' is a fixed closed subcomplex of K.

(11.1) $c^{n+1}(f) \subset K - K'$.

(11.2) $c^{n+1}(f) = 0$ *if and only if there is an extension* $f' \in Y^{K'+K^{n+1}}$ *of* f.

In fact, for each $\sigma^{n+1} \subset K'$ we have $f \mid \bar{\sigma}^{n+1} \simeq 0$, whence $c(f, \sigma^{n+1}) = 0$. Therefore (11.1) holds. (11.2) is a consequence of (9.2).

Let $f_0, f_1 \in Y^{K'+K^n}$, $f_0 \mid K' + K^{n-1} = f_1 \mid K' + K^{n-1}$.

(11.3) $d^n(f_0, f_1) \subset K - K'$.

(11.4) $d^n(f_0, f_1) = 0$ *if and only if* $f_0 \simeq f_1$ *rel.* $K' + K^{n-1}$.

For each $\sigma^n \subset K'$ we have $f_0 \mid \bar{\sigma}^n = f_1 \mid \bar{\sigma}^n$ and from (9.5) we deduce $d(f_0, f_1, \sigma^n) = 0$. Therefore (11.3) holds. (11.4) follows from (9.5).

(11.5) *Given* $f_0 \in Y^{K'+K^n}$ *and an* n*-chain* $d^n \subset K - K'$, *there is an* $f_1 \in Y^{K'+K^n}$ *such that* $f_0 \mid K' + K^{n-1} = f_1 \mid K' + K^{n-1}$ *and that* $d^n(f_0, f_1) = d^n$.

(11.6) *Given* $f \in Y^{K'+K^n}$ *and a cocycle* $c^{n+1} \subset K - K'$ *such that* $c^{n+1} \backsim c^{n+1}(f)$ *in* $K - K'$, *there is an* $f' \in Y^{K'+K^n}$ *such that* $f \mid K' + K^{n-1} = f' \mid K + K^{n-1}$ *and that* $c^{n+1}(f') = c^{n+1}$.

PROOFS. (11.5) is an immediate consequence of (8.4). In order to prove (11.6) let us consider an n-chain $d^n \subset K - K'$ such that

$$\delta d^n = c^{n+1}(f) - c^{n+1}.$$

By (11.5) there is an $f' \in Y^{K'+K^n}$ such that $f \mid K' + K^{n-1} = f' \mid K' + K^{n-1}$ and that $d^n(f, f') = d^n$. According to (10.2) we then have

$$\delta d^n = c^{n+1}(f) - c^{n+1}(f'),$$

and $c^{n+1}(f') = c^{n+1}$ follows.

12. EXTENSION THEOREM I. *Let* $f \in Y^{K'+K^n}$. *The* $(n+1)$*-chain* $c^{n+1}(f)$ *defined by* (9.1) *is a cocycle in* $K - K'$. *Moreover,* $c^{n+1}(f) \backsim 0$ *in* $K - K'$ *if and only if there is an* $f' \in Y^{K'+K^{n+1}}$ *such that* $f \mid K' + K^{n-1} = f' \mid K' + K^{n-1}$.

PROOF. The first part of the theorem follows from (10.1) and (11.1). If $c^{n+1}(f) \backsim 0$ in $K - K'$ then applying (11.6) for $c^{n+1} = 0$ we find an $f'' \in Y^{K'+K^n}$ such that $f \mid K' + K^{n-1} = f'' \mid K' + K^{n-1}$ and that $c^{n+1}(f'') = 0$. By (11.2) there is an extension $f' \in Y^{K'+K^{n+1}}$ of f''. We then have $f \mid K' + K^{n-1} = f' \mid K + K^{n-1}$. On the other hand, if such an f' exists we have: $\delta d^n(f, f') =$

$c^{n+1}(f) - c^{n+1}(f')$ because of (10.2), $c^{n+1}(f') = 0$ because of (11.2), and $d^n(f, f') \subset K - K'$ because of (11.3). It follows that $c^{n+1}(f) \backsim 0$ in $K - K'$.

HOMOTOPY THEOREM I. *Given two functions* f_0, $f_1 \in Y^K$ *such that* $f_0 \mid K' + K^{n-1} = f_1 \mid K' + K^{n-1}$, *the* n*-chain* $d^n(f_0, f_1)$ *defined by* (9.4) *is a cocycle in* $K - K'$. *Moreover,* $d^n(f_0, f_1) \backsim 0$ *in* $K - K'$ *if and only if*

$$f_0 \mid K' + K^n \simeq f_1 \mid K' + K^n \text{ rel. } K' + K^{n-2}. \tag{12.1}$$

PROOF. We shall consider the interval E as a complex containing two 0-cells 0 and 1 and one 1-cell ϵ oriented so that $\partial\epsilon = 0 - 1$. The product $L = K \times E$ is considered as a complex with cells of the form $\sigma^n \times 0$, $\sigma^n \times 1$ and $\sigma^n \times \epsilon$. The n-cell $\sigma^n \times 0$ is oriented as σ^n, $\sigma^n \times 1$ as $-\sigma^n$. The $(n+1)$-cell $\sigma^n \times \epsilon$ is oriented so as to have $\partial(\sigma^n \times \epsilon) = \partial\sigma^n \times \epsilon + \sigma^n \times \partial\epsilon$.

Let $L' = K' \times E + K \times 0 + K \times 1$. Let $g \in Y^{L'+L^n}$ be the map defined by

$$g(x, t) = f_0(x) \quad \text{for} \quad x \in K' + K^{n-1}, \qquad t \in E,$$

$$g(x, i) = f_i(x) \quad \text{for} \quad x \in K, \qquad i = 0, 1.$$

By (9.2) we then have $c(g, \sigma^{n+1} \times 0) = 0$ and $c(g, \sigma^{n+1} \times 1) = 0$. It follows from **8** that $c(g, \sigma^n \times \epsilon) = d(f_0, f_1, \sigma^n)$. Therefore we have

$$c^{n+1}(g) = d^n(f_0, f_1) \times \epsilon.$$

The condition $d^n(f_0, f_1) \backsim 0$ in $K - K'$ is therefore equivalent with the condition $c^{n+1}(g) \backsim 0$ in $L - L'$. By Ext. th. I this is equivalent with the existence of a $g' \in Y^{L'+L^{n+1}}$ such that $g \mid L' + L^{n-1} = g' \mid L' + L^{n-1}$. This, however, means exactly (12.1).

HOMOTOPY THEOREM IA. *If* ${}^n\boldsymbol{H}_{\pi_n}(K - K') = 0$ *and* $f_0, f_1 \in Y^K$ *then*

$$f_0 \mid K' + K^{n-1} \simeq f_1 \mid K' + K^{n-1} \text{ rel. } A \tag{12.2}$$

implies

$$f_0 \mid K' + K^n \simeq f_1 \mid K' + K^n \text{ rel. } A \tag{12.3}$$

for any subset A *of* K.

PROOF. By (5.2) and (12.2) there is an $f_0' \in Y^K$ such that $f_0' \mid K' + K^{n-1} = f_1 \mid K' + K^{n-1}$ and that $f_0 \simeq f_0'$ rel. A. Since $d^n(f_0', f_1) \backsim 0$ in $K - K'$ it follows from Hom. th. I that $f_0' \mid K' + K^n \simeq f_1 \mid K' + K^n$ rel. A. This implies (12.3).

13. In this section Y is i-simple for $i = n, n+1, \cdots, \dim(K - K')$.

EXTENSION THEOREM II. *If* ${}^{i+1}\boldsymbol{H}_{\pi_i}(K - K') = 0$ *for* $i = n, n+1, \cdots,$ *where* $\pi_i = \pi_i(Y)$, *then every* $f \in Y^{K'}$ *which has an extension* $f' \in Y^{K'+K^n}$ *has also an extension* $f'' \in Y^K$.

PROOF. We define a sequence $f_n = f', f_{n+1}, f_{n+2}, \cdots$ of maps $f_i \in Y^{K'+K^i}$ such that $f_{i+1} \mid K' + K^{i-1} = f_i \mid K' + K^{i-1}$. The sequence exists by Ext. th. I. Writing $f''(x) = \lim f_i(x)$ we have $f'' \in Y^K$ and $f \mid K' = f' \mid K' = f'' \mid K$.

HOMOTOPY THEOREM II. *If* ${}^i\boldsymbol{H}_{\pi_i}(K - K') = 0$ *for* $i = n, n+1, \cdots,$ *where* $\pi_i = \pi_i(Y)$, *then* $f_0, f_1 \in Y^K$ *and* $f_0 \mid K' + K^{n-1} \simeq f_1 \mid K' + K^{n-1}$ *rel.* A *imply* $f_0 \simeq f_1$ *rel.* A *for any subset* A *of* K'.

PROOF. Let $L = K \times E$, $L' = K' \times E + K \times 0 + K \times 1$. We then have ${}^{i+1}H_{\pi_i}(L - L') = 0$ for $i = n, n+1, \cdots$. It follows by Ext. th. II that every $g \in Y^L$ which has an extension $g' \in Y^{L'+L^n}$ also has an extension $g'' \in Y^L$.

If $f_0 \mid K' + K^{n-1} \simeq f_1 \mid K' + K^{n-1}$ rel. A, there is a $g' \in Y^{L'+L^n}$ such that

$$g'(x, t) = f_0(x) \quad \text{for} \quad x \in A, \qquad t \in E,$$

$$g'(x, i) = f_i(x) \quad \text{for} \quad x \in K, \qquad i = 0, 1.$$

Writing $g' \mid L = g$, we see that there is an extension $g'' \in Y^L$ of g. This implies $f_0 \simeq f_1$ rel. A.

IV. THE CASE $\pi_i(Y) = 0$ FOR $i < n$.

14. In this part we assume that $\pi_i(Y) = 0$ for $i < n$. If $n > 1$ this implies $\pi_1 = 0$ and by (7.3) Y is i-simple for all i. If $n = 1$ we assume that Y is 1-simple, or (see (7.1)) that π_1 is abelian. In particular everything can be applied for $Y = S^n$.

(14.1) *Every $f \in Y^{K'}$ has an extension $f' \in Y^{K'+K^n}$.*

(14.2) *f_0, $f_1 \in Y^K$ and $f_0 \mid K' \simeq f_1 \mid K'$ rel. A imply $f_0 \mid K' + K^{n-1} \simeq f_1 \mid K' + K^{n-1}$ rel. A for any $A \subset K'$.*

PROOFS. Obviously there is an extension $f_0 \in Y^{K'+K^0}$ of f. Since Y is arcwise connected (see **7**) there is also an extension $f_1 \in Y^{K'+K^1}$ of f. Since $\pi_i = 0$ for $i < n$, we have ${}^{i+1}H_{\pi_i}(K - K') = 0$ and applying Ext. th. I we obtain successively extensions $f_{i+1} \in Y^{K'+K^{i+1}}$ for $i = 1, 2, \cdots, n-1$.

Y being arcwise connected, it follows from $f_0 \mid K' \simeq f_1 \mid K'$ rel. A that $f_0 \mid K' + K^0 \simeq f_1 \mid K' + K^0$ rel. A. Since ${}^iH_{\pi_i}(K - K') = 0$ for $i = 1, 2, \cdots, n-1$, we obtain $f_0 \mid K' + K^{n-1} \simeq f_1 \mid K' + K^{n-1}$ rel. A applying Hom. th. I A.

15.

(15.1) *$f_0, f_1 \in Y^{K'+K^n}$ and $f_0 \mid K' \simeq f_1 \mid K'$ imply $c^{n+1}(f_0) \simeq c^{n+1}(f_1)$ in $K - K'$.*

PROOF. By (14.2) we have $f_0 \mid K' + K^{n-1} \simeq f_1 \mid K' + K^{n-1}$, and according to (5.2) there is an $f_0' \in Y^{K'+K^n}$ such that $f_0 \simeq f_0'$ and $f_0' \mid K' + K^{n-1} = f_1 \mid K' + K^{n-1}$. From (9.3), (10.2), and (11.3) we then have: $c^{n+1}(f_0) = c^{n+1}(f_0')$, $\delta d^n(f_0, f_1) = c^{n+1}(f_0') - c^{n+1}(f_1)$, and $d^n(f_0, f_1) \subset K^n - K'$. It follows that $c^{n+1}(f_0) \backsim c^{n+1}(f_1)$ in $K - K'$.

By (14.1) every $f \in Y^{K'}$ has an extension $f' \in Y^{K'+K^n}$. Let $\mathbf{c}^{n+1}(f)$ be the element of ${}^{n+1}H_{\pi_n}(K - K')$ determined by the cocycle $c^{n+1}(f')$. It follows from (15.1) that the choice of f' does not matter and that

(15.2) $\qquad f_0, f_1 \in Y^{K'}$ *and* $f_0 \simeq f_1$ *imply* $\mathbf{c}^{n+1}(f_0) = \mathbf{c}^{n+1}(f_1)$.

From Ext. th. I we obtain the following

EXTENSION THEOREM III. *Given $f \in Y^{K'}$ we have $\mathbf{c}^{n+1}(f) = 0$ if and only if there is an extension $f' \in Y^{K'+K^{n+1}}$ of f.*

16.

(16.1) *Given* $f_0, f_1, f_0', f_1' \in Y^K$ *such that* $f_0 \mid K' + K^{n-1} = f_1 \mid K' + K^{n-1}$, $f_0' \mid K' + K^{n-1} = f_1' \mid K' + K^{n-1}$, $f_0 \simeq f_0'$ *rel.* K' *and* $f_1 \simeq f_1'$ *rel.* K' *we have* $d^n(f_0, f_1) \frown d^n(f_0', f_1')$ *in* $K - K'$.

PROOF. Let $L = K \times E, L' = K' \times E + K \times 0 + K \times 1$. Let $g \in Y^{L'+L^n}$ be defined by

$$g(x, t) = f_0(x) \quad \text{for} \quad x \in K' + K^{n-1}, \qquad t \in E,$$

$$g(x, i) = f_i(x) \quad \text{for} \quad x \in K, \qquad i = 0, 1.$$

In an analogous way we define $g' \in Y^{L'+L^n}$ using f_0' and f_1' instead of f_0 and f_1. As in **12** we then have

$$c^{n+1}(g) = d^n(f_0, f_1) \times \epsilon, \qquad c^{n+1}(g') = d^n(f_0', f_1') \times \epsilon. \tag{16.2}$$

Now, $f_0 \simeq f_0'$ rel. K' and $f_1 \simeq f_1'$ rel. K' obviously imply $g \mid L' \simeq g' \mid L'$, and therefore, by (15.1), $c^{n+1}(g) \frown c^{n+1}(g')$ in $L - L'$. Using (16.2) we then obtain $d^n(f_0, f_1) \frown d^n(f_0, f_1)$ in $K - K'$.

Given two functions $f_0, f_1, \in Y^K$ such that $f_0 \mid K' = f_1 \mid K'$ we have, by (14.2), $f_0 \mid K' + K^{n-1} \simeq f_1 \mid K' + K^{n-1}$ rel. K', and by (5.2) there are two functions $f_0', f_1' \in Y^K$, such that

$$f_0 \simeq f_0' \text{ rel. } K', \; f_1 \simeq f_1' \text{ rel. } K', \qquad f_0' \mid K' + K^{n-1} = f_1' \mid K' + K^{n-1}. \tag{16.3}$$

Let $\boldsymbol{d}^n(f_0, f_1)$ be the element of ${}^n\boldsymbol{H}_{\pi_n}(K - K')$ determined by the cocycle $d^n(f_0', f_1')$. It follows from (16.1) that $\boldsymbol{d}^n(f_0, f_1)$ is independent of the particular choice of f_0' and f_1' such that (16.3) holds. By (10.3) and (10.4) we have

(16.4) *If* $f_0, f_1, f_2 \in Y^K$ *and* $f_0 \mid K' = f_1 \mid K' = f_2 \mid K'$ *then* $\boldsymbol{d}^n(f_0, f_1) + \boldsymbol{d}^n(f_1, f_2) = \boldsymbol{d}^n(f_0, f_2)$,

$$\boldsymbol{d}^n(f_0, f_1) = -\boldsymbol{d}^n(f_1, f_0). \tag{16.5}$$

HOMOTOPY THEOREM III. *Given* $f_0, f_1 \in Y^K$ *such that* $f_0 \mid K' = f_1 \mid K'$ *we have* $\boldsymbol{d}^n(f_0, f_1) = 0$ *if and only if*

$$f_0 \mid K' + K^n \simeq f_1 \mid K' + K^n \text{ rel. } K'. \tag{16.6}$$

PROOF. If $\boldsymbol{d}^n(f_0, f_1) = 0$ then $d^n(f_0', f_1') \frown 0$ in $K - K'$ and by Hom. th. I we have $f_0' \mid K' + K^n \simeq f_1' \mid K' + K^n$ rel. K'. Using (16.3) we obtain (16.6). On the other hand, by (5.2), (16.6) implies the existence of an $f_0' \in Y^K$ such that $f_0 \simeq f_0'$ rel. K' and that $f_0' \mid K' + K^n = f_1 \mid K' + K^n$. It follows $d^n(f_0', f_1) = 0$ and therefore $\boldsymbol{d}^n(f_0, f_1) = 0$.

17. In this section (and in **18**) we assume as before that $\pi_i(Y) = 0$ for $i < n$. If $n = 1$ we assume that Y is i-simple for $i = 1, 2, \cdots, \dim (K - K')$. As before, everything can be applied for $Y = S^n$.

Combining Ext. th. II and III we have the

EXTENSION THEOREM IV. *Let* ${}^{i+1}\boldsymbol{H}_{\pi_i}(K - K') = 0$ *for* $i = n + 1, n + 2, \cdots,$

where $\pi_i = \pi_i(Y)$. *Given* $f \in Y^{K'}$ *we have* $\mathbf{c}^{n+1}(f) = 0$ *if and only if there is an extension* $f' \in Y^K$ *of* f.

Similarly Hom. th. II and III imply

HOMOTOPY THEOREM IV. *Let* ${}^i\boldsymbol{H}_{\pi_i}(K - K') = 0$ *for* $i = n+1, n+2, \cdots$, *where* $\pi_i = \pi_i(Y)$. *Given* $f_0, f_1 \in Y^K$ *such that* $f_0 \mid K' = f_1 \mid K'$ *we have* $\mathbf{d}^n(f_0, f_1) = 0$ *if and only if* $f_0 \simeq f_1$ *rel.* K'.

The hypotheses of Ext. th. IV and of Hom. th. IV are obviously satisfied if $K - K' \subset K^n$.

EXISTENCE THEOREM. *Let* ${}^{i+1}\boldsymbol{H}_{\pi_i}(K - K') = 0$ *for* $i = n+1, n+2, \cdots$, *where* $\pi_i = \pi_i(Y)$. *Given* $f_1 \in Y^K$ *and* $\mathbf{d}^n \in {}^n\boldsymbol{H}_{\pi_n}(K - K')$ *there is an* $f_0 \in Y^K$ *such that* $f_0 \mid K' = f_1 \mid K'$ *and that* $\mathbf{d}^n(f_0, f_1) = \mathbf{d}^n$.

PROOF. Let d^n be a cocycle of the cohomology class $\mathbf{d}^n$. By (11.5) and (10.4) there is a map $f_0 \in Y^{K' + K^n}$ such that $f_0 \mid K' + K^{n-1} = f_1 \mid K' + K^{n-1}$ and that $d^n(f_0, f_1) = d^n$. By (10.2) we have $\delta d^n = c^{n+1}(f_0) - c^{n+1}(f_1)$. Since $\delta d^n = 0$ and $c^{n+1}(f_1) = 0$ (see (11.2)) it follows $c^{n+1}(f_0) = 0$ and by (11.2) there is an extension $f_0' \in Y^{K' + K^{n+1}}$ of f_0.

Now, let $K'' = K' + K^n$. We then have ${}^{i+1}\boldsymbol{H}_{\pi_i}(K - K'') = 0$ for $i = n+1, n+2, \cdots$ and by Ext. th. II there is an extension $f_0'' \in Y^K$ of f_0. We then have $d^n(f_0'', f_1) = d^n$ and therefore $\mathbf{d}^n(f_0'', f_1) = \mathbf{d}^n$.

18. Let $f^* \in Y^K$. As in **4** $Y^K(K', f^*)$ will be the sub-family of Y^K containing all $f \in Y^K$ such that $f \mid K' = f^* \mid K'$. The family $Y^K(K', f^*)$ is divided into homotopy classes rel. K', two maps $f_0, f_1, \in Y^K(K, f^*)$ being in the same class if, and only if, $f_0 \simeq f_1$ rel. K'.

Given a homotopy class Φ rel. K' of $Y^K(K', f^*)$ we define $\mathbf{d}^n(\Phi) = \mathbf{d}^n(f, f^*)$ where $f \in \Phi$. It follows from (16.4), (16.5), and Hom. th. III, that the element $\mathbf{d}^n(\Phi)$ of ${}^n\boldsymbol{H}_{\pi_n}(K - K')$ is defined uniquely. Under the hypothesis of Hom. th. IV we have $\Phi_0 = \Phi_1$ if and only if $\mathbf{d}^n(\Phi_0) = \mathbf{d}^n(\Phi_1)$. Under the hypothesis of the Existence th. there is for each $\mathbf{d}^n \in {}^n\boldsymbol{H}_{\pi_n}(K - K')$ a homotopy class Φ rel. K' of $Y^K(K', f^*)$ such that $\mathbf{d}^n(\Phi) = \mathbf{d}^n$. We obtain, therefore,

CLASSIFICATION THEOREM I. *Let* ${}^i\boldsymbol{H}_{\pi_i}(K - K') = {}^{i+1}\boldsymbol{H}_{\pi_i}(K - K') = 0$ *for* $i = n+1, n+2, \cdots$, *where* $\pi_i = \pi_i(Y)$. *The elements of* ${}^n\boldsymbol{H}_{\pi_n}(K - K')$ *are in a* $(1-1)$*-correspondence with the homotopy classes rel.* K' *of* $Y^K(K', f^*)$. *The correspondence is determined by the operation* $\mathbf{d}^n(f, f^*)$.

Taking $K' = 0$ and $f^* = 0$ we obtain

CLASSIFICATION THEOREM II. *Let* ${}^i\boldsymbol{H}_{\pi_i}(K) = {}^{i+1}\boldsymbol{H}_{\pi_i}(K) = 0$ *for* $i = n+1, n+2, \cdots$ *where* $\pi_i = \pi_i(Y)$. *The elements of* ${}^n\boldsymbol{H}_{\pi_n}(K)$ *are in a* $(1-1)$*-correspondence with the homotopy classes of* Y^K. *The correspondence is determined by the operation* $\mathbf{d}^n(f) (= \mathbf{d}^n(f, 0))$.

Note that according to **16** $\mathbf{d}^n(f)$ is defined as follows: by (14.2) we have $f \mid K^{n-1} \simeq 0$, therefore by (5.2) there is an $f' \in Y^K$ such that $f \simeq f'$ and that $f' \mid K^{n-1} = 0$. $\mathbf{d}^n(f)$ is then the element of ${}^n\boldsymbol{H}_{\pi_n}(K)$ corresponding to the cocycle $d^n(f') = d^n(f', 0)$.

The hypothesis of Class. th. I are obviously satisfied if $K - K' \subset K^n$. Tak-

ing $K = K^n$ in Class. th. II we obtain the theorem of Whitney quoted in the introduction.

V. Application to Homology

19. Let G, H, Z be three abelian groups. If to each $\alpha \in G$ and $\beta \in H$ there corresponds and $\alpha \cdot \beta$ in Z, and both distributive laws are satisfied, we say[18] G and H form a *group-pair* with respect to Z. If I is the group of rational integers then clearly G and I form a group pair with respect to G.

Given a finite n-chain $A^n = \sum_i \alpha_i \sigma_i^n$ with $\alpha_i \in G$ and an arbitrary n-chain $B^n = \sum_i \beta_i \sigma_i^n$ with $\beta_i \in H$ we write

$$A^n \cdot B^n = \sum_i \alpha_i \cdot \beta_i .$$

It is easy to see that

$$\partial A^{n+1} \cdot B^n = A^{n+1} \cdot \delta B^n$$

for every finite $(n+1)$-chain A^{n+1}. It follows that if A_0^n and A_1^n are finite cycles mod K' (coef. G) and B_0^n and B_1^n are cocycles in $K - K'$ (coef. H) then

$$A_0^n \sim^* A_1^n \quad \text{and} \quad B_0^n \backsim B_1^n$$

imply

$$A_0^n \cdot B_0^n = A_1^n \cdot B_1^n .$$

We see then that ${}^n\boldsymbol{H}^{*G}(K)$ mod K' and ${}^n\boldsymbol{H}_H(K - K')$ form a group pair with respect to Z. Similar relations hold for ${}^n\boldsymbol{H}^G(K)$ mod K' and ${}^n\boldsymbol{H}_H^*(K - K')$.

20. We assume that

1°) *Y is locally connected in dimensions $\leqq n$*[19]

2°) *$\pi_i(Y) = 0$ for $i < n$*

3°) *$\pi_n(Y)$ is isomorphic with I.*

It follows from (7.3) and (7.1) that Y is n-simple.

It follows from our hypothesis that $\pi_n(Y)$ can be considered as identical with the n^{th} homology group ${}^n\mathcal{H}^1(Y)$ with integer coefficients.[20] Further, there is in Y an n-dimensional cycle Γ_0^n (coef. I) such that for every n-dimensional cycle Γ^n (coef. G) in Y there is a unique $\alpha \in G$ such that $\Gamma^n \sim \alpha\Gamma_0^n$.[21] Owing to this fact we may write ${}^n\mathcal{H}^G(Y) = G$ and in particular $\pi_n(Y) = {}^n\mathcal{H}^I(Y) = I$.

21. Let $f \in Y^{K'}$. For every finite cycle A^n in K' with coefficients in G we have

$$f(A^n) \sim \alpha\Gamma_0^n \quad \text{where} \quad \alpha \in G.$$

---

[18] See E. Čech, Ann. of Math. 37(1936), p. 684.

[19] See e.g. C. Kuratowski, Fund. Math. 24(1935), p. 269.

[20] W. Hurewicz, Proc. Akad. Amsterdam 38(1935), pp. 521–2. Explanations concerning homology in Y will be found there.

[21] See N. E. Steenrod, Amer. Jour. of Math. 58(1936), pp. 661–701.

The element α will be called the *degree of f on A^n* and denoted by $g(f, A^n)$. It is obvious that $A_0^n \sim^* A_1^n$ in K' implies $g(f, A_0^n) = g(f, A_1^n)$ and therefore $g(f, \mathfrak{a}^n)$ is defined for every $\mathfrak{a}^n \in {}^nH^{*G}(K')$ and is a homomorphic mapping of ${}^nH^{*G}(K')$ into G. Obviously $f_0, f_1 \in Y^{K'}$ and $f_0 \simeq f_1$ imply $g(f_0, \mathfrak{a}^n) = g(f_1, \mathfrak{a}^n)$.

Let $f' \in Y^{K'+K^n}$ be an extension of f. For any $(n+1)$-cell σ_i^{n+1} we then have according to **9** and **20** $c(f', \sigma_i^{n+1}) = g(f', \partial\sigma_i^{n+1})$ which can be written

$$\sigma_i^{n+1} \cdot c^{n+1}(f') = g(f', \partial\sigma_i^{n+1}).$$

It follows that

$$A^{n+1} \cdot c^{n+1}(f') = g(f', \partial A^{n+1})$$

for any finite $(n+1)$-chain A^{n+1} with coefficients in G. If, in particular, A^{n+1} is a finite cycle mod K' then $\partial A^{n+1} \subset K'$ and

$$A^{n+1} \cdot c^{n+1}(f') = g(f, \partial A^{n+1}). \tag{21.1}$$

Therefore, according to **15** and **19**

$$\mathfrak{a}^{n+1} \cdot c^{n+1}(f) = g(f, \partial\mathfrak{a}^{n+1}) \text{ for every } \mathfrak{a}^{n+1} \in {}^{n+1}H^{*G}(K) \bmod K'. \tag{21.2}$$

22. Let $f_0, f_1 \in Y^K$ and $f_0 \mid K' = f_1 \mid K'$. For every finite cycle A^n mod K' (coef. G) $f_0(A^n) - f_1(A^n)$ is an n-cycle in Y, and therefore $f_0(A^n) - f_1(A^n) \sim \alpha\Gamma_0^n$ for some $\alpha \in G$. We write $g(f_0, f_1, A^n) = \alpha$. As before we verify easily that $g(f_0, f_1, \mathfrak{a}^n) \in G$ is a homomorphic map defined for every $\mathfrak{a}^n \in {}^nH^{*G}(K) \bmod K'$. Clearly $f_0', f_1' \in Y^K, f_0 \simeq f_0'$ rel. K' and $f_1 \simeq f_1'$ rel. K' imply $g(f_0, f_1, \mathfrak{a}^n) = (f_0', f_1', \mathfrak{a}^n)$.

Now let us suppose that $f_0' \mid K' + K^{n-1} = f_1' \mid K' + K^{n-1}$. We then have according to **8**, **9**, and **20** $d(f_0', f_1', \sigma_i^n) = g(f_0', f_1', \sigma_i^n)$ and as before

$$A^n \cdot d^n(f_0', f_1') = g(f_0', f_1', A^n)$$

and therefore

$$A^n \cdot d^n(f_0', f_1') = g(f_0, f_1, A^n) \tag{22.1}$$

for every finite cycle A^n mod K'. According to **16** and **19** we then have

$$\mathfrak{a}^n \cdot d^n(f_0, f_1) = g(f_0, f_1, \mathfrak{a}^n) \text{ for every } \mathfrak{a}^n \in {}^nH^{*G}(K) \bmod K'. \tag{22.2}$$

23. We assume now (besides the hypothesis on Y made in **20**) that

1°) $K - K'$ *is finite,*

2°) $G = R$ *is the group of real numbers reduced* mod 1.

In this case the groups ${}^nH^R(K) \bmod K'$[22] and ${}^nH_I(K - K')$ are orthogonal[23] and therefore every element of ${}^nH_I(K - K')$ can be considered as a character[24] of

---

[22] This group has to be considered as a topological compact group. See L. Pontrjagin, Ann. of Math. 35(1934), p. 908.

[23] L. Pontrjagin, Ann. of Math. 35(1934), pp. 361-388; H. Whitney, Duke Math. Jour. 3(1937), p. 40.

[24] L. Pontrjagin, Loc. cit.

${}^nH^R(K)$ mod K'. By (21.2) and (22.2) we see that the cohomology classes $c^{n+1}(f) \in {}^{n+1}H_I(K - K')$ and $d^n(f_0, f_1) \in {}^nH_I(K - K')$, when considered as characters, are just equal to $g(f, \partial a^{n+1})$ and $g(f_0, f_1, a^n)$.[25] Using Ext. th. III and Hom. th. III we obtain therefore

EXTENSION THEOREM III*. *Given $f \in Y^{K'}$ we have $g(f, \partial a^{n+1}) = 0$ for every $a^{n+1} \in {}^{n+1}H^R(K)$ mod K' if and only if there is an extension $f' \in Y^{K'+K^{n+1}}$ of f.*

HOMOTOPY THEOREM III*. *Given $f_0, f_1 \in Y^K$ such that $f_0 \mid K' = f_1 \mid K'$ we have $g(f_0, f_1, a^n) = 0$ for every $a^n \in {}^nH^R(K)$ mod K' if and only if $f_0 \mid K' + K^n \simeq f_1 \mid K' + K^n$ rel. K'.*

24. In order to obtain classification theorems in terms of homology we have to admit that

1°) *Y is locally connected in dimensions $\leqq \dim (K - K')$;*

2°) *$\pi_i(Y) = 0$ for $i < n$;*

3°) *$\pi_n(Y)$ is isomorphic with I.*

If $n > 1$ it follows from (7.3) that Y is i-simple for all i. If $n = 1$ we require Y to be i-simple for $i = 1, 2, \cdots \dim (K - K')$. In particular we may take $Y = S^n$.

The group $\pi_i = \pi_i(Y)$ being countable at most for $i \leqq \dim (K - K')$, there is[26] a topological compact group ρ_i orthogonal to π_i. In particular we may take $\rho_n = R$.

If we admit further that $K - K'$ is finite, then the groups ${}^jH^{\rho_i}(K)$ mod K'[22] and ${}^jH_{\pi_i}(K - K')$ are orthogonal[23] and therefore the formulas

$$ {}^jH^{\rho_i}(K) \text{ mod } K' = 0, \qquad {}^jH_{\pi_i}(K - K') = 0 $$

are equivalent.

Using the argument of **23** we may restate all the theorems of **17** and **18** replacing cohomology by homology. In particular we obtain

CLASSIFICATION THEOREM I*. *Let ${}^iH^{\rho_i}(K)$ mod $K' = {}^{i+1}H^{\rho_i}(K)$ mod $K' = 0$ for $i = n + 1, n + 2, \cdots$. The characters of the group ${}^nH^R(K)$ mod K' are in a $(1 - 1)$-correspondence with the homotopy classes rel. K' of $Y^K(K', f^*)$. To each $f \in Y^K(K', f^*)$ there corresponds the character $g(f, f^*, a^n)$.*

CLASSIFICATION THEOREM II*.[27] *Let ${}^iH^{\rho_i}(K) = {}^{i+1}H^{\rho_i}(K) = 0$ for $i = n + 1, n + 2, \cdots$. The characters of the group ${}^nH^R(K)$ are in a $(1 - 1)$-correspondence with the homotopy classes of Y^K. To each $f \in Y^K$ there corresponds the character $g(f, a^n)$.*

VI. MANIFOLDS

25. Let Q^r be a finite or infinite geometrical cell-complex which is an oriented r-dimensional combinatorial manifold.[28] The first barycentric subdivision

[25] This conclusion (and consequently also the theorems which follow) can be obtained even when $K - K'$ is infinite provided the groups ${}^iH^R(K - K')$ and ${}^iH_I^*(K)$ mod K' are orthogonal for $i = 1, 2, \cdots$.

[26] L. Pontrjagin, Ann. of Math. 35(1934), pp. 361–388.

[27] See S. Eilenberg, Compositio Math. 6(1939), p. 429.

[28] See e.g. K. Reidemeister, *Topologie der Polyeder*, Leipzig 1938, p. 151.

Q_1^r of Q^r is a simplicial complex, and using the simplices of Q_1^r we may define as usual the (oriented) dual $(r-n)$-cell for each n-cell of Q^r. The dual cells form a cell-complex K^r which is also an r-manifold and has Q_1^r as a barycentric subdivision.

If σ_i^{r-n} is the $(r-n)$-cell of K^r dual to the n-cell τ_1^n of Q^r we write $\mathfrak{D}(\tau_i^n) = \sigma_i^{r-n}$ and $\mathfrak{D}^*(\sigma_i^{r-n}) = \tau_i^n$. More generally, given an n-chain.

$$A^n = \sum_i \alpha_i \tau_i^n$$

we write

$$\mathfrak{D}(A^n) = \sum_i \alpha_i \sigma_i^{r-n}$$

and $\mathfrak{D}^*[\mathfrak{D}(A^n)] = A^n$. It is well known that

$$\partial \mathfrak{D}(A^n) = (-1)^{n+1} \mathfrak{D}(\delta A^n).^{29}$$

Similar relations hold for $\mathfrak{D}^*$. We obtain thus $(1-1)$-isomorphisms $\mathfrak{D}[{}^nH^G(Q^r)] = {}^{r-n}H_G(K^r)$ and $\mathfrak{D}[{}^nH_G(Q^r)] = {}^{r-n}H^G(K^r)$. The inverse isomorphisms are given by $\mathfrak{D}^*$.

26. P will stand for an arbitrary closed subcomplex of Q^r. We shall denote by $\mathfrak{D}(P)$ the subcomplex of K^r consisting of all the cells $\mathfrak{D}(\tau_i^n)$ where $\tau_i^n \in P$.

(26.1) $P \subset \mathfrak{D}(P)$.

(26.2) *$P_1 \subset P_2$ implies $\mathfrak{D}(P_1) \subset \mathfrak{D}(P_2)$.*

(26.3) *$K^r - \mathfrak{D}(P)$ is a closed subcomplex of K^r.*

(26.4) *$P = P^i$ implies $K^{r-i-1} \subset K^r - \mathfrak{D}(P)$.*

(26.5) *$\mathfrak{D}(A^n) \subset \mathfrak{D}(P)$ for every n-chain A^n in P. $\mathfrak{D}^*(A^n) \subset P$ for every n-chain A^n in $\mathfrak{D}(P)$.*

(26.6) *$K^r - \mathfrak{D}(P)$ is a deformation retract[30] of $Q^r - P$.*[31]

(26.7) *Every $f \in Y^{K^r - \mathfrak{D}(P)}$ has an extension $f' \in Y^{Q^r - P}$.*

(26.8) *$f_0, f_1 \in Y^{Q^r - P}$ and $f_0 \mid K^r - \mathfrak{D}(P) \simeq f_1 \mid K^r - \mathfrak{D}(P)$ imply $f_0 \simeq f_1$.*

(26.1)–(26.5) follow directly from the definition. In order to prove (26.6), notice that $K^r - \mathfrak{D}(P)$ consists of all simplices of the barycentric subdivision Q_1^r of Q^r which have no vertex on P. (26.6) is therefore a consequence of the following quite general and elementary lemma:

Let Q_1 be a geometric simplicial complex, P_1 a closed subcomplex of Q_1 and $C(P_1)$ the closed subcomplex of Q_1 consisting of all simplices of Q_1 which have no vertex on P_1. The complex $C(P_1)$ is then a deformation retract of $Q_1 - P_1$.

(26.7) and (26.8) follow from (26.6).

27. We assume that Y is n-simple. All chains will have coefficients from $\pi_n = \pi_n(Y)$.

---

[29] See H. Seifert and W. Threfall, *Lehrbuch der Topologie*, Leipzig-Berlin 1934, p. 245.

[30] $A \subset X$ is a deformation retract of X if there is a map $g \in X^{X \times E}$ such that $g(x, 0) = x$, $g(x, 1) \in A$ for $x \in X$ and $g(x, 1) = x$ for $x \in A$.

[31] See S. Lefschetz, *Topology*, New York 1930, p. 141.

Let P be a closed subcomplex of Q^r and let $f \in Y^{Q^r - P^{r-n-1}}$. Taking $K' = K^r - \mathfrak{D}(P)$ we have according to (26.1) and (26.4)

$$K' + K^n \subset K^r - \mathfrak{D}(P^{r-n-1}) \subset Q^r - P^{r-n-1}.$$

Write

$$f_1 = f \mid K' + K^n, \qquad \gamma^{r-n-1}(f) = \mathfrak{D}^*[c^{n+1}(f_1)].^{32}$$

(27.1) *$\gamma^{r-n-1}(f)$ is a cycle in P^{r-n-1} (with coef. in $\pi_n(Y)$).*

(27.2) *$f' \in Y^{Q^r - P^{r-n-1}}$ and $f \simeq f'$ imply $\gamma^{r-n-1}(f) = \gamma^{r-n-1}(f')$.*

(27.3) *If $\gamma^{r-n-1}(f) \subset P_1 \subset P$ then there is an $f' \in Y^{Q^r - (P_1 + P^{r-n-2})}$ such that $f \simeq f' \mid Q^r - P^{r-n-1}$.*

(27.4) *$\gamma^{r-n-1}(f) = 0$ if and only if there is an $f' \in Y^{Q^r - P^{r-n-2}}$ such that $f \simeq f' \mid Q^r - P^{r-n-1}$.*

(27.5) *Given an $(r-n-1)$-cycle $\gamma^{r-n-1} \subset P$ such that $\gamma^{r-n-1} \sim \gamma^{r-n-1}(f)$ in P, there is an $f' \in Y^{Q^r - P^{r-n-1}}$ such that $\gamma^{r-n-1}(f') = \gamma^{r-n-1}$ and that $f \mid Q^r - P^{r-n} \simeq f' \mid Q^r - P^{r-n}$.*

(27.6) *$\gamma^{r-n-1}(f) \sim 0$ in P if and only if there is an $f' \in Y^{Q^r - P^{r-n-2}}$ such that $f \mid Q^r - P^{r-n} \simeq f' \mid Q^r - P^{r-n}$.*

Proofs. Ad (27.1). By (10.1) and (11.1) $c^{n+1}(f_1)$ is a cocycle in $\mathfrak{D}(P)$. Therefore, by (26.5), $\gamma^{r-n-1}(f)$ is a cycle in P.

Ad (27.2). $f \simeq f'$ implies $f \mid K' + K^n \simeq f' \mid K + K^n$. Therefore $f_1 \simeq f_1'$, and by (9.3) $c^{n+1}(f_1) = c^{n+1}(f_1)$. It follows that $\gamma^{r-n-1}(f) = \gamma^{r-n-1}(f')$.

Ad (27.3). $\gamma^{r-n-1}(f) \subset P_1$ implies $c^{n+1}(f_1) \subset \mathfrak{D}(P_1)$, whence $c(f_1, \sigma_i^{n+1}) = 0$ for every $\sigma_i^{n+1} \subset \mathfrak{D}(P) - \mathfrak{D}(P_1)$. By (9.2) f_1 admits an extension on every such σ_i^{n+1} and therefore there is an extension $f_1' \in Y^{K^r - \mathfrak{D}(P_1 - P^{r-n-2})}$ of f_1. According to (26.7) there is an extension $f' \in Y^{Q^r - (P_1 - P^{r-n-2})}$ of f_1'. We then have $f \mid K^r - \mathfrak{D}(P^{r-n-1}) = f_1 = f' \mid K^r - \mathfrak{D}(P^{r-n-1})$ and by (26.8) $f \simeq f' \mid Q^r - P^{r-n-1}$.

Ad (27.4). If $\gamma^{r-n-1}(f) = 0$ then taking $P_1 = 0$ in (27.3) we obtain an $f' \in Y^{Q^r - P^{r-n-2}}$ such that $f \simeq f' \mid Q^r - P^{r-n-1}$. On the other hand if such an f' exists we have $\gamma^{r-n-1}(f) = \gamma^{r-n-1}(f')$ by (27.2) and therefore $\gamma^{r-n-1}(f) = 0$ by (27.1) since $\gamma^{r-n-1}(f') \subset P^{r-n-2}$.

Ad (27.5). Let $c^{n+1} = \mathfrak{D}(\gamma^{r-n-1})$. Then $\gamma^{r-n-1} \sim \gamma^{r-n-1}(f)$ in P implies $\mathfrak{D}(\gamma^{r-n-1}) \backsim \mathfrak{D}[\gamma^{r-n-1}(f)]$ in $\mathfrak{D}(P)$ and therefore $c^{n+1} \backsim c^{n+1}(f_1)$ in $K^r - K'$. By (11.6) there is an $f_1' \in Y^{K' + K^n}$ such that $f_1 \mid K' + K^{n-1} = f_1' \mid K' + K^{n-1}$ and that $c^{n+1}(f_1') = c^{n+1}$. Since $K' + K^n = K^r - \mathfrak{D}(P^{r-n-1})$ there is, by (26.7), an extension $f' \in Y^{Q^r - P^{r-n-1}}$ of f_1'. We then have $\gamma^{r-n-1}(f') = \mathfrak{D}[c^{n+1}(f_1')] = \gamma^{r-n-1}$. Since $K' + K^{n-1} = K^r - \mathfrak{D}(P^{r-n})$ and $f \mid K' + K^{n-1} = f_1 \mid K + K^{n-1} = f_1' \mid K' + K^{n-1} = f' \mid K + K^{n-1}$, we have $f \mid K^r - \mathfrak{D}(P^{r-n}) = f' \mid K^r - \mathfrak{D}(P^{r-n})$, and by (26.8) $f \mid Q^r - P^{r-n} \simeq f' \mid Q^r - P^{r-n}$.

Ad (27.6). If $\gamma^{r-n-1}(f) \sim 0$ in P, then, taking $\gamma^{r-n-1} = 0$ in (27.5), we obtain an $f'' \in Y^{Q^r - P^{r-n-1}}$ such that $\gamma^{r-n-1}(f'') = 0$ and that $f \mid Q^r - P^{r-n} \simeq$

---

[32] This definition of $\gamma^{r-n-1}(f)$ is obviously equivalent with that given in the introduction. Cf. footnote [15].

$f'' \mid Q^r - P^{r-n}$. Applying (27.4) we obtain an $f' \in Y^{Q^r - P^{r-n-2}}$ such that $f \mid Q^r - P^{r-n} \simeq f' \mid Q^r - P^{r-n}$. On the other hand, if such an f' exists then $f \mid K' + K^{n-1} \simeq f' \mid K' + K^{n-1}$ since $K' + K^{n-1} \subset Q^r - P^{r-n}$. By (5.2) we may therefore suppose that $f \mid K' + K^{n-1} = f' \mid K' + K^{n-1}$. Since $K' + K^n \subset Q^r - P^{r-n-1}$ and $K' + K^{n+1} \subset Q^r - P^{r-n-2}$ it follows from Ext. th. I that $c^{n+1}(f \mid K' + K^n) \backsim 0$ in $K - K' = \mathfrak{D}(P)$ and therefore that $\gamma^{r-n-1}(f) \sim 0$ in P.

28. In this section we assume that $\pi_i(Y) = 0$ for $i < n$ if $n > 1$, and that Y is 1-simple (i.e. that $\pi_1(Y)$ is abelian) if $n = 1$.

(28.1) *Given $f \in Y^{Q^r - P}$ there is an $f' \in Y^{Q^r - P^{r-n-1}}$ such that $f \simeq f' \mid Q^r - P$.*

(28.2) *$f_0, f_1 \in Y^{Q^r - P^{r-n}}$ and $f_0 \mid Q^r - P \simeq f_1 \mid Q^r - P$ imply $f_0 \simeq f_1$.*

(28.3) *$f_0, f_1 \in Y^{Q^r - P^{r-n-1}}$ and $f_0 \mid Q^r - P \simeq f_1 \mid Q^r - P$ imply $\gamma^{r-n-1}(f_0) \sim \gamma^{r-n-1}(f_1)$ in P.*

Proofs. Ad (28.1). By (14.1) there is an $f_1 \in Y^{K' + K^n}$ such that $f \mid K' = f_1 \mid K'$, where as before we take $K' = K^r - \mathfrak{D}(P)$. Since $K' + K^n = K^r - \mathfrak{D}(P^{r-n-1})$ there is, by (26.7), an extension $f' \in Y^{Q^r - P^{r-n-1}}$ of f_1. We then have $f \mid K^r - \mathfrak{D}(P) = f' \mid K^r - \mathfrak{D}(P)$ and therefore $f \simeq f' \mid Q^r - P$ by (26.8).

Ad (28.2). $f_0 \mid Q^r - P \simeq f_1 \mid Q^r - P$ implies $f_0 \mid K' \simeq f_1 \mid K'$ and by (14.2) $f_0 \mid K' + K^{n-1} \simeq f_1 \mid K' + K^{n-1}$. Since $K' + K^{n-1} = K^r - \mathfrak{D}(P^{r-n})$ it follows from (26.8) that $f_0 \simeq f_1$.

Ad (28.3). $f_0 \mid Q^r - P \simeq f_1 \mid Q^r - P$ implies $f_0 \mid K' \simeq f_1 \mid K'$, and by (15.1) $c^{n+1}(f_0 \mid K' + K^n) \backsim c^{n+1}(f_1 \mid K' + K^n)$ in $K^r - K' = \mathfrak{D}(P)$. It follows that $\gamma^{r-n-1}(f_0) \sim \gamma^{r-n-1}(f_1)$ in P.

Let $f \in Y^{Q^r - P}$ and let f' be given by (28.1). Let $\gamma^{r-n-1}(f)$ be the element of ${}^{r-n-1}H^{\pi_n}(P)$ determined by the cycle $\gamma^{r-n-1}(f')$. It follows from (28.3) that $\gamma^{r-n-1}(f)$ is independent of the choice of f'.

(28.4) $\gamma^{r-n-1}(f) = \mathfrak{D}^*[c^{n+1}(f \mid K')]$ *where* $K' = K^r - \mathfrak{D}(P)$.

(28.5) *Given $f \in Y^{Q^r - P}$ we have $\gamma^{r-n-1}(f) = 0$ if and only if there is an $f' \in Y^{Q^r - P^{r-n-2}}$ such that $f \simeq f' \mid Q^r - P$.*

Proofs. (28.4) follows straight from the definition of γ^{r-n-1} and c^{n+1}. If $\gamma^{r-n-1}(f) = 0$ then for every $f'' \in Y^{Q^r - P^{r-n-1}}$ such that $f \simeq f'' \mid Q^r - P$ we have $\gamma^{r-n-1}(f'') \sim 0$ in P. It follows by (27.6) that there is an $f' \in Y^{Q^r - P^{r-n-2}}$ such that $f \simeq f'' \mid Q^r - P \simeq f' \mid Q^r - P$. On the other hand, if such an f' exists we have $\gamma^{r-n-1}(f') = 0$ by (27.4), and therefore $\gamma^{r-n-1}(f) = 0$.

29. We assume now that $\pi_i(Y) = 0$ for $i < n$ if $n > 1$ and that Y is i-simple for $i = 1, 2, \cdots, r-1$ if $n = 1$.

(29.1) *Let ${}^{r-i-1}H^{\pi_i}(P) = 0$ for $i = n+1, n+2, \cdots, r-1$ where $\pi_i = \pi_i(Y)$. Given $f \in Y^{Q^r - P}$ we have $\gamma^{r-n-1}(f) = 0$ if and only if there is an $f' \in Y^{Q^r}$ such that $f \simeq f' \mid Q^r - P$.*

This follows from (28.5) and (27.6) applied successively for $n+1, n+2, \cdots, r-1$.

30. Let G and H be two groups forming a group pair with respect to a group Z. Given a finite n-chain $A^n = \sum_i \alpha_i \sigma_i^n$ in K^r with coefficients in G and an arbitrary $(r-n)$-chain $B^{r-n} = \sum_i \beta_i \tau_i^{r-n}$ in Q^r with coefficients in H, we write

$$\chi(A^n, A^{r-n}) = \sum_i \alpha_i \cdot \beta_i .^{33}$$

We obviously have

$$\chi(A^n, A^{r-n}) = A^n \cdot \mathfrak{D}(B^{r-n}) = \mathfrak{D}^*(A^n) \cdot B^{r-n}. \tag{30.1}$$

It follows therefore from **19**, **25**, and **26** that ${}^n\boldsymbol{H}^{*G}(K^r)$ mod $K^r - \mathfrak{D}(P)$ and ${}^{r-n}\boldsymbol{H}^H(P)$ form a group pair with respect to Z. Similar relations hold for ${}^n\boldsymbol{H}^G(K^r)$ mod $K^r - \mathfrak{D}(P)$ and ${}^{r-n}\boldsymbol{H}^{*H}(P)$.

31. We make the same hypothesis about Y as in **20**. Let $f \in Y^{Q^{r-P}}$. By (21.2), (28.4) and (30.1) we have

(31.1) $\chi[\boldsymbol{a}^{n+1}, \boldsymbol{\gamma}^{r-n-1}(f)] = g(f, \partial \boldsymbol{a}^{n+1})$ *for every* $\boldsymbol{a}^{n+1} \in {}^{n+1}\boldsymbol{H}^{*G}(K^r)$ mod $K^r - \mathfrak{D}(P)$

Now, if P is finite and $G = R$ (see **23** and footnote[25]) the complex $\mathfrak{D}(P)$ is finite and ${}^{n+1}\boldsymbol{H}^R(K^r)$ mod $K^r - \mathfrak{D}(P)$ is orthogonal to ${}^{r-n-1}\boldsymbol{H}^I(P)$, since it is orthogonal to ${}^{n-r-1}\boldsymbol{H}_I[\mathfrak{D}(P)]$ (see **23**). It follows therefore from (31.1) that $\boldsymbol{\gamma}^{n-r-1}(f)$ considered as a character of ${}^{n+1}\boldsymbol{H}^R(K^r)$ mod $K^r - \mathfrak{D}(P)$ is equal to $g(f, \partial \boldsymbol{a}^{n+1})$. Therefore it follows from (28.5) that

(31.2) *Given* $f \in Y^{Q^{r-P}}$ *we have* $g(f, \partial \boldsymbol{a}^{n+1}) = 0$ *for every* $\boldsymbol{a}^{n+1} \in {}^{n+1}\boldsymbol{H}^R(K^r)$ mod $K^r - \mathfrak{D}(P)$ *if and only if there is an* $f' \in Y^{Q^{r-P^{r-n-2}}}$ *such that* $f \simeq f' \mid Q^r - P$.

Appendix I. On Normal Mappings[33]

Let Y be an arbitrary topological space and let $y_0 \in Y$. A map $f \in Y^{K^n}$ will be called n-normal if $f \mid K^{n-1} = 0$ (i.e. if $f(x) = y_0$ for $x \in K^{n-1}$).

Given an oriented n-cell σ_i^n in K^n, the map $f \mid \bar{\sigma}_i^n$ defines (if f is n-normal) uniquely an element $d(f, \sigma_i^n)$ of $\pi_n(Y)$. If $n > 1$ then $\pi_n(Y)$ is abelian and we may define the n-chain $d^n(f)$ and the $(n+1)$-cocycle $c^{n+1}(f)$ by

$$d^n(f) = \sum_i d(f, \sigma_i^n)\sigma_i^n , \qquad c^{n+1}(f) = \delta d^n(f).$$

$c(f, \sigma_i^{n+1})$ will be defined as the coefficient of the $(n+1)$-cell σ_i^{n+1} in $c^{n+1}(f)$. Given two n-normal maps $f_0 , f_1 \in Y^K$ we take

$$d^n(f_0 , f_1) = d^n(f_0) - d^n(f_1).$$

Starting from the definition of $\pi_n(Y)$ we can prove that $c(f, \sigma_i^{n+1}) = 0$ if and only if f can be extended on σ_i^{n+1}. It follows that

$d^n(f)$ *is a cocycle* (i.e. $c^{n+1}(f) = 0$) *if and only if there is an extension* $f' \in Y^{K^{n+1}}$ *of* f.

---

[33] The purpose of this appendix is to make clear the position of the results of H. Whitney (Duke Math. Jour. 3(1937), pp. 51–55) in the theory developed in this paper. The proofs are on the same lines as those of Whitney and may be left to the reader.

From this, statements analogous to those of **9–12** can easily be deduced. In particular we have

EXTENSION THEOREM. *Given an n-normal map $f \in Y^{K'}$, the n-chain $d^n(f) \subset K'$ is part of a cocycle in K[34] if and only if there is an n-normal extension $f' \in Y^{K'+K^{n+1}}$ of f.*

HOMOTOPY THEOREM. *Given two n-normal maps f_0, $f_1 \in Y^K$ such that $f_0 \mid K' = f_1 \mid K'$, the n-chain $d^n(f_0, f_1)$ is a cocycle in $K - K'$. Moreover $d^n(f_0, f_1) \frown 0$ in $K - K'$ if and only if $f_0 \mid K' + K^n \simeq f_1 \mid K' + K^n$ rel. $K' + K^{n-2}$.*

APPENDIX II. MAPPINGS OF INFINITE $(n+1)$-MANIFOLDS

We assume that $\pi_i(Y) = 0$ for $i < n$ if $n > 1$ and that Y is i-simple for $i = 1, 2$ if $n = 1$. In particular we may take $Y = S^n$.

Let Q^{n+1} be an *infinite* geometrical cell-complex which is an oriented $(n+1)$-dimensional combinatorial manifold.

CLASSIFICATION THEOREM. *The homotopy classes of $Y^{Q^{n+1}}$ are in (1-1)-correspondence with the elements of the group ${}^1\boldsymbol{H}^{\pi_n}(Q^{n+1})$ where $\pi_n = \pi_n(Y)$.*

PROOF. Let K^{n+1} be the dual of Q^{n+1}. Since Q^{n+1} is a connected infinite complex, it is easy to see that ${}^0\boldsymbol{H}^G(Q^{n+1}) = 0$ and therefore, by **25**, that ${}^{n+1}\boldsymbol{H}_G(K^{n+1}) = 0$ for every abelian group G. The hypotheses of Class. th. II are thus satisfied and the homotopy classes of $Y^{K^{n+1}} = Y^{Q^{n+1}}$ are in a $(1-1)$-correspondence with the elements of ${}^n\boldsymbol{H}_{\pi_n}(K^{n+1})$. By **25** this group is isomorphic with ${}^1\boldsymbol{H}^{\pi_n}(Q^{n+1})$. This proves the theorem.

UNIVERSITY OF MICHIGAN

[34] Whitney, loc. cit., p. 53. It is easy to verify that $d^n(f)$ is part of a cocycle in K if and only if $c^{n+1}(f) \frown 0$ in $K - K'$; ibid., p. 54.

ANNALS OF MATHEMATICS
Vol. 41, No. 1, January, 1940

WEAK TOPOLOGIES OF NORMED LINEAR SPACES

BY LEON ALAOGLU

(Received January 21, 1939)

Introduction

The concepts of weak convergence introduced by Banach [9, pp. 122, 133][1] are of exceeding interest, but in certain respects are rather restricted. Many of the results concerned with these concepts are in general valid only for separable spaces. It is natural, then, to turn to the general theory of limits of Moore and Smith [2] for a generalization. A unified theory of weak convergence is presented in Section 1, together with some of its consequences for weak convergence of functionals. Of these the principal result is Theorem 1:3, which states that the unit sphere of the adjoint of a Banach space is bicompact in the weak topology of functionals. This is a generalization of a result given by Banach [9, p. 123] which is valid only for separable spaces.

Characterizations of spaces equivalent or isomorphic to adjoint spaces are given in Section 2, and are applied to a class of spaces considered by Dunford and Morse [13]. A result on universal spaces is given in Section 3.

The remainder of the paper is devoted to the development of theories of integration and differentiation of abstract functions. Two theories of integration are outlined in Section 4, both of which can be obtained from that of Birkhoff [11] by replacing unconditional convergence of series by its analogues for weak convergence of functionals and elements. The second of these integrals can be specialized from that of Gelfand [15] by the addition of a uniformity condition.

In Section 5 a theory of differentiation of functions with values in an adjoint space is developed. Instead of defining the derivative as the limit of the difference quotients of the function, the class of all ultimate limit points of the difference quotients in the sense of weak convergence of functionals is considered. For a function f of bounded absolute variation in the sense of Definition 5:1, it is shown that these sets are non-null almost everywhere, so that they define a multiple-valued function, the "derivative" of f. By using the first of the integrals defined in Section 4, it is shown that the fundamental theorem of the calculus holds for "derivatives" of absolute continuous functions. Gelfand's result on weak derivatives [15], and the results of Clarkson [12], Dunford and Morse [13], Gelfand [15], and Pettis [18] on strong derivatives are derived from this result. It is of interest to note that in the derivation of Clarkson's theorem, it is shown that uniformly convex spaces are regular.

[1] The numbers in brackets refer to the bibliography.

These theories of integration and differentiation are finally applied to obtain two representation theorems, one for linear functionals on the space of Bochner integrable functions, and the other for linear transformations on the space of Lebesgue integrable functions to an arbitrary Banach space.

1. Weak convergence

It has been pointed out that in the theory of normed linear spaces there arise two concepts of weak convergence different from convergence in the norm. If the space E^* is adjoint to a Banach space E, a sequence (x_n) of elements of E is said to have the *weak limit* x if $\lim_n f(x_n) = f(x)$ for every element f of E^* [9, p. 133]. Similarly, a sequence (f_n) of elements of E^* is said to be *weakly convergent* if $\lim_n f_n(x)$ exists for every element x of E [9, p. 122]. In adjoint spaces these two types of convergence are not always identical.

Banach has used these two types of convergence notably in the study of what he has called regular closure. His results in this connection are valid only for separable spaces. To remove these separability conditions, we make use of the general theory of limits of Moore and Smith [2].

The central concept of this theory is that of a *directed set.* This name is given to a class A of elements α, together with a binary relation R between the elements of A, if the relation R is transitive, and if further, for every pair α_1, α_2 of elements of A, there is a third element α such that the relations $\alpha R\alpha_1$, $\alpha R\alpha_2$, hold. Limits of functions (b_α) on A to the real numbers are defined in the following manner.

DEFINITION 1:1.

(1) $\limsup_\alpha b_\alpha = \inf_\alpha (\sup_{\alpha_1 R\alpha} b_{\alpha_1})$;

(2) $\liminf_\alpha b_\alpha = -\limsup_\alpha (-b_\alpha)$;

(3) $b = \lim_\alpha b_\alpha$ *if for every* $\epsilon > 0$ *there is an* α_ϵ *such that* $|b - b_\alpha| < \epsilon$ *if* $\alpha R\alpha_\epsilon$ *holds.*

These limits are well-defined, and $\lim_\alpha b_\alpha$ is a linear functional; that is, $\lim_\alpha (aa_\alpha + bb_\alpha) = a \lim_\alpha a_\alpha + b \lim_\alpha b_\alpha$. We shall use the term "directed system of elements of E" for a function on a directed set A to a Banach space E, and we shall not specify the directed set whenever it is arbitrary.

DEFINITION 1:2. *If G is a total subset of the space E^* adjoint to E, and if (x_α) is a directed system of elements of E, we shall say that the system is weakly convergent on G if* $\sup_\alpha \| x_\alpha \|$ *is finite, and if* $\lim_\alpha f(x_\alpha)$ *exists for every f of G. If for some element x of E, $f(x) = \lim_\alpha f(x_\alpha)$ for every element f of G, we shall write*

$$x = \text{w-}\lim_\alpha x_\alpha \text{ on } G.$$

If G is E^*, and if the system (x_n) is a sequence, this reduces to Banach's definition of weak convergence of elements. For weak convergence of functionals, the subset G of E^{**} can be taken to be the set E_0 of all functionals $\phi_x(f) = f(x)$, where x is in E.

Another example is due to Schauder [6, p. 61]. In a space E with a basis (ξ_n) let G be the set of coefficient functionals in the development $x = \sum_n a_n(x)\xi_n$.

The sequences weakly convergent on G are then weakly convergent in Schauder's sense.

The condition that G be total in Definition 1:2 is natural; for otherwise limits would not be unique. The condition that the directed system be bounded is superfluous if the system is a sequence, and if G is E^*; for arbitrary directed systems and arbitrary G this need not be so. For, if E is infinite-dimensional, a discontinuous, real-valued linear functional ϕ on E can be constructed by the use of a Hamel base [9, p. 231] for E. Let A be the class of all finite-dimensional linear subspaces α of E. Then A with the inclusion relation $\supset$ is a directed set. But the section of ϕ on α is continuous, and hence is the section on α of a linear functional f_α continuous on E. Thus $\phi(x) = \lim_\alpha f_\alpha(x)$ for every x; but the system (f_α) is not bounded in view of the next theorem.

THEOREM 1:1. *If the directed system (f_α) of elements of E^* is weakly convergent on E, and if $f(x) = \lim_\alpha f_\alpha(x)$, then f is linear and continuous on E, and $\| f \| \leqq \liminf_\alpha \| f_\alpha \|$.*

The linearity of f is a consequence of the linearity of the limit operation; as for the continuity of f, we see that $| f(x) | = \lim_\alpha | f_\alpha(x) | \leqq \liminf_\alpha \| f_\alpha \| \cdot \| x \|$.

The weak convergence on G of Definition 1:2 defines a topology in E, in which the bounded subspaces of E are completely regular. Let K be a bounded subset of E. Then if σ is a finite subset of G, $\epsilon > 0$, and x an element of K, let

$$(1{:}1) \qquad U_{\sigma\epsilon}(x) = [\text{all } x_1 \text{ in } K \text{ such that } | g(x - x_1) | < \epsilon \text{ for every } g \text{ in } \sigma].$$

THEOREM 1:2. *The bounded subspace K of E with neighborhood system* (1:1) *is a completely regular Hausdorff space of character not exceeding $\aleph_0$ times the cardinal number of G. This topology is equivalent to the topology of K defined by the directed systems of elements of K weakly convergent on G to limits in K. Equivalent neighborhood systems are obtained if G is replaced by any set having the same linear closed extension as G.*

By the *character* of a space is meant the least cardinal number of a basis of the open sets in the space.

It is readily verified that G can be replaced by its linear extension G_L in Definition 1:2 without affecting the character of the convergence of any system. The proof of the statement that G can be replaced by its metric closure G_C is similar to the proof of Banach's Theorem 1 [9, p. 133].

Let the element x of K be the weak limit on G of the directed system (x_α) of elements of a subset K_0 of K. Then for every g_i in a finite subset σ of G, and for every $\epsilon > 0$, there is an α_i such that $| g_i(x - x_\alpha) | < \epsilon$ if $\alpha R \alpha_i$ holds. But there is an α_ϵ such that $\alpha R \alpha_i$ holds for every i, and then x_α is in $U_{\sigma\epsilon}(x)$ if $\alpha R \alpha_\epsilon$ holds. Conversely, if every neighborhood $U_{\sigma\epsilon}(x)$ of the element x of K contains an element $x_{(\sigma\epsilon)}$ of K_0, $x = \text{w-lim}_{(\sigma\epsilon)} x_{(\sigma\epsilon)}$, where $(\sigma\epsilon) R (\sigma_1\epsilon_1)$ means that $U_{\sigma\epsilon}(x)$ is contained in $U_{\sigma_1\epsilon_1}(x)$.

We can now replace G by the set G_0 of all functionals of the form $g/M \| g \|$, with g in G; and M can be chosen so that $| g_0(x) | \leqq 1$ if x is in K and g_0 in G_0.

Now consider the class Q of all functions ϕ on G_0 to the closed interval $(-1, 1)$. The class Q is a bicompact Hausdorff space with the neighborhoods

(1:2) $\quad U_{\sigma\epsilon}(\phi) = [\text{all } \phi_1 \text{ in } Q \text{ such that } |\phi(g) - \phi_1(g)| < \epsilon \text{ for every } g \text{ of } \sigma],$

where σ is a finite subset of G_0, $\epsilon > 0$, and ϕ is an element of Q. This space has been studied by Tychonoff [8, p. 546], who proves these results, and shows that the character of Q is $\aleph_0$ times the cardinal number of G_0. But, on comparing the neighborhoods (1:1) with those of (1:2), we see that K is a subspace of Q, and hence is completely regular, as is shown by Tychonoff [8, p. 545]. Also, the cardinal number of G_0 does not exceed that of G.

THEOREM 1:3. *The unit sphere K of the space E^* adjoint to E, with the neighborhoods* (1:1) *for $G = E$, is a bicompact Hausdorff space. Its character is $\aleph_0$ times the least cardinal number τ of a subset of E total for E^*.*

Banach has shown that the linear closed extension of a subset H of E total for E^* is the whole space E [9, p. 58]. Hence, by Theorem 1:2, E can be replaced by a total subset G_0 of the unit sphere of E in defining the neighborhoods (1:1). K is then a subspace of the bicompact space Q of Tychonoff. Moreover, K is closed in Q, so that K is bicompact. For, let the element f of Q be in the closure of the subspace K. Then every neighborhood $U_{\sigma\epsilon}(f)$ contains an element $f_{\sigma\epsilon}$ of K, and as in the proof of Theorem 1:2, $f(x) = \lim_{\sigma\epsilon} f_{\sigma\epsilon}(x)$ for every x of G_0. But then by Theorem 1:2, the system is weakly convergent, and by Theorem 1:1, it has a limit in K.

Since G_0 can be chosen to have minimal cardinal number τ, the character χ of K does not exceed $\tau\aleph_0$. The local character χ_1 of K at $f = 0$ does not exceed χ; and Alexandroff [5, pp. 269–70] has shown that from the complete system $U_{\sigma\epsilon}(0)$ of neighborhoods of $f = 0$ a complete subsystem $V_{\sigma\epsilon}$ can be chosen of minimal cardinal number χ_1. The character χ_1 is infinite, so that we can extend the range of ϵ in the $V_{\sigma\epsilon}$ to include all positive rational numbers. Let H_1 be the logical sum of all the sets σ occurring in the $V_{\sigma\epsilon}$. Then, since χ_1 is infinite, the cardinal number of H_1 does not exceed χ_1. Furthermore, H_1 is total. For, if the element f of K vanishes on H_1 the sequence $f_n = f$ converges weakly both to f and to zero, so that $f = 0$. Hence $\chi = \tau\aleph_0$, since $\tau\aleph_0 \leqq \chi_1 \leqq \chi \leqq \tau\aleph_0$.

When E is separable, $\tau \leqq \aleph_0$, and the space K satisfies the second countability axiom. Theorem 1:3 has been obtained with this restriction by Banach [9, pp. 123–4].

A similar restriction occurs in Banach's treatment of the regular closure. We recall here the definitions of regular closure and transfinite closure [9, pp. 115 ff.].

DEFINITION 1:3. (1) *If θ is a limit ordinal, the transfinite sequence $(f_\xi \mid \xi < \theta)$ of elements of E^* is said to have the transfinite limit f if the system (f_ξ) is bounded, and if for every element x of E,*

$$\liminf_\xi f_\xi(x) \leqq f(x) \leqq \limsup_\xi f_\xi(x).$$

(2) *If every bounded transfinite sequence of elements of a subset G of E^* has a transfinite limit in G, G is said to be transfinitely closed.*

(3) *The subset G of E^* is regularly closed if for every functional f not in G there is an element x of E such that $f(x) \neq 0$, while $g(x) = 0$ for every element g of G.*

(4) *The subset G of E^* is said to be weakly closed on E if every directed system of elements of G converging weakly on E has a limit in G.*

Banach [9, pp. 121, 125] has shown that these three closure properties are equivalent for linear subspaces G of E^* if E is separable. In the next theorem we show that this equivalence holds in arbitrary Banach spaces.

THEOREM 1:4. *If the subspace G of E is linear, the following properties of G are equivalent:*

(1) *G is transfinitely closed,*

(2) *G is regularly closed,*

(3) *G is weakly closed on E.*

The first implication, that (1) implies (2), has been proved by Banach [9, p. 119]. To show that (2) implies (3), let f be the weak limit on E of a directed system of elements of G. Then, if $g(x)$ vanishes for every element g of G, so does $f(x)$, so that f is in G. Finally, to prove that (3) implies (1), let (f_ξ) be a bounded transfinite sequence of elements of G. Then the sets

$$G_\eta = [\text{all } f_\xi \text{ with } \eta \leqq \xi < \theta]$$

lie in a metrically closed sphere K, which is bicompact in the weak topology of Theorem 1:3. Hence the intersection of the weak closures of the sets G_η is non-null [4, p. 259]. Any element of the intersection is in G, and is a transfinite limit of the system (f_ξ).

THEOREM 1:5. *If the functional F is linear and continuous on E^*, and if whenever the system (f_α) of elements of E^* converges weakly to f on E, it is true that $F(f) = \lim_\alpha F(f_\alpha)$, then there is an x_0 in E such that $F(f) = f(x_0)$ for every f of E^*. The converse is also true.*

The proof of this theorem is the same as that of Banach's Theorem 8 [9, p. 131], which was proved for separable spaces.

2. Two characterizations of adjoint spaces

THEOREM 2:1. *If E is a Banach space, the following two conditions on E are equivalent:*

(1) *There exists a Banach space E_1 such that E is isomorphic* [9, p. 180] *to E_1^*;*

(2) *There is a total subset G of E^* and a constant $M > 0$ such that for every element x of E, $\| x \| \leqq M \sup_{f \text{ in } G_L} | f(x) | / \| f \|$, where G_L is the linear extension of G; and every directed system (x_α) of elements of E weakly convergent on G has a limit in E.*

If (2) holds, for every linear and continuous functional F on the linear closed extension G_{LC} of G there is an element x_0 of E such that $F(f) = f(x_0)$ for every f of G_{LC}, and such that $\| x_0 \| \leqq M \cdot \| F \|$.

If E is isomorphic to E_1^*, let T be the linear transformation on E to E_1^* which

is effective. Then $\| T^{-1} \|^{-1} \| x \| \leqq \| Tx \| \leqq \| T \| \cdot \| x \|$. Let G be the set of all functionals $f_y(x) = (Tx)(y)$ with y in E_1. This set is total, since $(Tx)(y) = 0$ for every y if and only if $Tx = 0$, which is equivalent to $x = 0$. To obtain the inequality involving the norm in E, we observe that $| f_y(x) | = | (Tx)(y) | \leqq \| y \| \cdot \| T \| \cdot \| x \|$, so that $\| f_y \| \leqq \| T \| \cdot \| y \|$. Hence we see that

$$\sup_y | f_y(x) | / \| f_y \| \geqq \sup_y | (Tx)(y) | / \| T \| \cdot \| y \| = \| Tx \| / \| T \| \geqq \| x \| / [\| T \| \cdot \| T^{-1} \|].$$

The system (x_α) converges weakly on G if and only if the system (Tx_α) converges on E_1. But, by Theorem 1:1, this latter system has a limit g in E_1^*, and $T^{-1}g = \text{w-lim}_\alpha x_\alpha$ on G. This completes the proof of the necessity of (2).

If (2) holds for G, it also holds for the linear closed extension G_{LC} of G, by Theorem 1:2. The functionals $\phi_x(f) = f(x)$ are linear and continuous on G_{LC} and the totality of these functionals is a total linear subspace of $(G_{LC})^*$ which is weakly closed on G. Hence, by Theorem 1:4, it is the whole space $(G_{LC})^*$, so that every F in $(G_{LC})^*$ is a ϕ_x, and $\| x \| \leqq M \sup_{f \text{ in } G_{LC}} | \phi_x(f) | / \| f \| = M \| \phi_x \|$. This transformation carrying x into ϕ_x is the isomorphism between E and $(G_{LC})^*$.

COROLLARY 2:1. *If in* (1) *of Theorem* 2:1 *the word "isomorphic" is replaced by "equivalent"* [9, p. 180], (2) *may be replaced by*:

(2′) *There is a total subset G of E^* such that* $\| x \| = \sup_{f \text{ in } G_L} | f(x) | / \| f \|$, *and such that the unit sphere of E is bicompact in the topology of Theorem* 1:2 *defined by G.*

If (2′) *holds, E is equivalent to $(G_{LC})^*$ under the transformation indicated in Theorem* 2:1.

In this case, $1 = \| T \| = \| T^{-1} \|$, and from the construction of the set G, the bicompactness follows from Theorem 1:3. To prove that (2′) is sufficient, note that the transformation at the end of Theorem 2:1 is an equivalence in this case.

THEOREM 2:2. *If the Banach space E has a basis (ξ_n) such that the series $\sum_n a_n \xi_n$ converges in the norm whenever its partial sums are bounded, then E is isomorphic to $(G_{LC})^*$, where G is the set of coefficient functionals in the development $x = \sum_n a_n(x) \xi_n$.*

Spaces of this type have been considered by Dunford and Morse [13, p. 415], who showed that the space has an equivalent norm such that for every sequence (a_i), the sequence $(\| \sum_{i=1}^n a_i \xi_i \|)$ is non-decreasing. Since G is denumerable, it suffices to show that the unit sphere of E is weakly compact in the topology defined by G. If (x_n) is a bounded sequence, the sequences $(a_i(x_n) \mid n)$ are bounded for each i, so that, by the diagonal process, a subsequence (x_{n_k}) can be chosen such that $a_i = \lim_k a_i(x_{n_k})$ exists for every i. The series $\sum_i a_i \xi_i$ then converges to an element x of E, and $x = \text{w-lim}_k x_{n_k}$ on G. Also

$$\| x \| = \sup_n \| \sum_{i=1}^n a_i(x) \xi_i \| = \sup_{n,f} | \sum_{i=1}^n a_i(x) f(\xi_i) | / \| f \|.$$

But $\sum_{i=1}^{n} a_i f(\xi_i)$ is in G_L, and $\| \sum_{i=1}^{n} a_i f(\xi_i) \| \leqq \| f \|$, so that condition (2′) of Corollary 2:1 is satisfied.

3. Universal Banach Spaces

Theorem 1:3 suggests the possibility of constructing universal Banach spaces. Let K be the bicompact space of Theorem 1:3, and let $\phi_x = (f(x) \mid f \text{ in } K)$. Then ϕ_x is continuous on K, and is linear in x. If we define $\| \phi_x \|$ to be $\sup_{f \text{ in } K} | \phi_x(f) |$, we see that $\| x \| = \| \phi_x \|$, so that this transformation is an isometry between E and a subspace of the space of continuous functions on the bicompact space K.

When K satisfies the second countability axiom, it is well-known that K is a continuous image of Cantor's ternary set $\mathfrak{B}_{\aleph_0}$. But if the character χ of K exceeds $\aleph_0$, such a result is not known. Stone [14] has constructed the bicompact Boolean spaces $\mathfrak{B}_\tau$ of character τ, consisting of all functions b on a range P of cardinal number τ with values zero or unity. A basis of $\mathfrak{B}_\tau$ is formed by the finite intersections of the sets

$$U_p = [\text{all } b \text{ in } \mathfrak{B}_\tau \text{ such that } b(p) = 0],$$

$$U_p' = \mathfrak{B}_\tau - U_p .$$

Stone has shown that a bicompact Hausdorff space of character τ is a continuous image of a closed subset of $\mathfrak{B}_\tau$. Let T be the transformation effective for K. Then the functions $\psi_x(b) = \phi_x(Tb)$ are continuous on $\mathfrak{B}$, linear in x, and $\| x \| = \sup_{b \text{ in } \mathfrak{B}} | \psi_x(b) |$. This completes the proof of the following theorem.

THEOREM 3:1. *If the least cardinal number of a total subset of the adjoint E^* of the Banach space E is τ, then E is equivalent to a linear closed subspace of the space of all continuous real-valued functions f on a bicompact Boolean space $\mathfrak{B}$ of character $\tau\aleph_0$, where $\| f \| = \sup_{b \text{ in } \mathfrak{B}} | f(b) |$.*

When $\tau \leqq \aleph_0$, we can take $\mathfrak{B}$ to be $\mathfrak{B}_{\aleph_0}$, as was done by Banach [9, pp. 185–7]. But if $\tau > \aleph_0$, the space $\mathfrak{B}$ is not uniquely determined by τ. One might conjecture that Stone's theorem could be strengthened to read "every bicompact Hausdorff space of character τ is a continuous image of $\mathfrak{B}_\tau$," and then we would have a stronger Theorem 3:1.

4. Integration of abstract functions

Gelfand [15] has given a definition of an integral for abstract-valued functions T on a domain P with a measure function. If for every element f of the adjoint E^* of the range of T the real function $f[T(p)]$ is integrable in the sense of Lebesgue as defined by Fréchet [1], the function T is said to be weakly integrable. If the range of T is the adjoint E^* of a Banach space E, one might define T to be integrable if $\phi_x[T(p)]$ is integrable for every x, where, as before, $\phi_x(f) = f(x)$. It is here that one can make a construction similar to that of Birkhoff [11]. Birkhoff's construction makes use of the norm topology of the range of T; a

similar theory using weak convergence is a natural extension. First it is necessary to have a notion of unconditional weak convergence of series. This is at hand. For use in the following definition, we observe that the class of all finite sets σ of positive integers is a directed set with the inclusion relation $\sigma_1 \supset \sigma_2$.

DEFINITION 4:1. *A series $\sum_i f_i$ of elements of the space E^* adjoint to a Banach space E is said to be weakly unconditionally convergent on E if $\lim_\sigma \sum_{i \text{ in } \sigma} f_i(x)$ exists for every element x of E.*

This is the analogue for weak convergence on E of Orlicz' notion of unconditional convergence in the norm [21, p. 242]. A similar definition has been given by Gelfand [16, p. 246].

THEOREM 4:1. *A necessary and sufficient condition that the series $\sum_i f_i$ of elements of E^* be weakly unconditionally convergent on E is that*

$$\sup_\sigma \| \sum_{i \text{ in } \sigma} f_i \| < +\infty.$$

If this be so, w-$\lim_\sigma \sum_{i \text{ in } \sigma} f_i$ *exists on E.*

The class of all σ is denumerable, so that, if the condition of the definition is satisfied, the limit exists semi-uniformly. Hence, by Hildebrandt's Theorem [3, p. 313], there is a σ_1 and a constant $M > 0$ such that if σ contains σ_1, $\| \sum_{i \text{ in } \sigma} f_i \| < M$. Then for any σ, $\| \sum_{i \text{ in } \sigma} f_i \| < M + \sum_{i \text{ in } \sigma_1} \| f_i \|$ holds. Conversely, if $\sup_\sigma \| \sum_{i \text{ in } \sigma} f_i \| < +\infty$, the series $\sum_{i=1}^\infty | f_i(x) |$ is convergent for every element x of E, and $\lim_\sigma \sum_{i \text{ in } \sigma} f_i(x) = \sum_{i=1}^\infty f_i(x)$. It is then a consequence of Theorem 1:1 that w-$\lim_\sigma \sum_{i \text{ in } \sigma} f_i$ on E exists.

One can now proceed to replace unconditional convergence by weak unconditional convergence in Birkhoff's definitions and results on the summation of series of bounded subsets of E^*. At the same time the metric closure $\bar{B}$ must be replaced by the weak closure hB. This is defined to be the set of all limits of directed systems of elements of B weakly convergent on E. If B is a bounded subset of E^*, the set hB is weakly closed on E in the sense of Definition 1:3 (4). All of Birkhoff's results in this connection are valid when so amended.

The functions to be studied are multiple-valued functions on a domain P of the type described by Birkhoff to the adjoint E^* of a Banach space E. Such a function is said to be *weakly summable relative to the partition* $\Pi = (A_i)$ of P if the series of sets $\sum_i T(A_i)mA_i$ is weakly summable in the sense of weak convergence on E as just indicated. If T is weakly summable relative to Π, *the weak integral range of T relative to* Π is defined to be $J_\Pi(T) = h[Co(\sum_i T(A_i)mA_i)]$. The function is then said to be *integrable over P on E* if it is weakly summable relative to some partition Π, and if the integral ranges of T have only one point in common, which is denoted by $\int_P T(p)\, dm$. This integral is characterized in the following theorem.

THEOREM 4:2. *A function T on P to E^* is integrable over P on E if and only if it is summable relative to some partition* Π, *and if the real functions $\phi_x[T(p)]$ are integrable over P in the sense of Lebesgue.*

Thus this integral differs from the modification of Gelfand's integral indi-

cated at the beginning of this section by the additional condition of summability, which is a sort of uniformity condition. It can be shown that this condition of integrability is weaker than that of Birkhoff, and that when the latter holds, the integrals are equal.

Much use has been made of the transformation ψ on E to E^{**} defined by the relations $\psi(x) = \phi_x$, and $\phi_x(f) = f(x)$. By means of this transformation one other integral can be defined. A function T on P to the Banach space E is said to be *integrable over P on E^** if the function $\psi[T(p)]$ with values in E^{**} is integrable over P on E^*. By Theorem 4:2 this condition can be shown to be stronger than Gelfand's, and weaker than that of Birkhoff.

5. Differentiation of abstract functions

In this section we shall consider single-valued functions on a real interval to the adjoint E^* of a Banach space E, and we shall be concerned with a theory of differentiation for such functions. As in the Lebesgue theory, a notion of *bounded variation* is necessary.

DEFINITION 5:1. *A single-valued function on the interval (c, d) to the Banach space E is said to be of bounded absolute variation if there exists a number M such that for every partition $c = a_0 \leqq a_1 \leqq \cdots \leqq a_n = d$ of the interval, it is true that*

$$\sum_{i=1}^{n} \| f(a_i) - f(a_{i-1}) \| \leqq M.$$

There is associated with each type of convergence of infinite series in a Banach space E a notion of function of bounded variation. The word *absolute* used here refers to this distinction.

The derivative of a real-valued function is the limit $f'(t) = \lim_{a=0} [f(t + a) - f(t)]/a$. If $f'(t)$ exists, it is the only element of the set

$$\prod_{d>0} h[\text{all } [f(t + a) - f(t)]/a \text{ with } |a| \leqq d],$$

where h denotes the closure of the indicated class of numbers. Conversely, if this set contains only one element, $f'(t)$ exists and is that element.

DEFINITION 5:2. *Let $f(\cdot, t)$ be a single-valued function on the interval (c, d) to the adjoint E^* of a Banach space E. If for some t on the interval there exist constants M and d_1 such that $\| f(\cdot, t + a) - f(\cdot, t) \| < M |a|$ when $|a| < d_1$, the set*

$$\phi(\cdot, t) = \prod_{0<d\leqq d_1} h[\text{all } [f(\cdot, t + a) - f(\cdot, t)]/a \text{ with } |a| < d],$$

is defined to be the "derivative" of f at t. As before, the letter h denotes the weak closure on E of the set indicated.

If the condition of the definition is satisfied, the sets occurring therein are bounded and weakly closed on E. But these sets are bicompact in the topology of Theorem 1:3, and they form a monotone decreasing system. Hence their intersection is non-null, in consequence of a lemma on bicompact spaces which is well-known in a special case [4, p. 259].

LEMMA 5:1. *If (G_α) is a directed system of closed non-null subsets of a bicompact Hausdorff space K, and if further, $G_{\alpha_1} \subset G_{\alpha_2}$ whenever $\alpha_1 R \alpha_2$ holds, then $\Pi_\alpha G_\alpha$ is non-null.*

If the sets have no point in common, for every point p of K there is a set $G_\alpha(p)$ not containing p, and hence a neighborhood U_p of p such that $U_p G_\alpha(p) = 0$. These neighborhoods cover K; and since K is bicompact, a finite number U_{p_i} will cover the whole space. But then if α stands in the relation R to the α_i corresponding to the points p_i, $G_\alpha U_{p_i} = 0$ for all i. This is a contradiction.

THEOREM 5:1. *If the function $f(\cdot, t)$ on the interval (c, d) to the space E^* is of bounded absolute variation, it has a "derivative" $\phi(\cdot, t)$ almost everywhere, which is integrable over (c, d) on E. Except on a set of zero measure, $\phi(\cdot, t) = \psi(\cdot, t)g(t)$, where $\psi(\cdot, t)$ is bounded, and $g(t)$ is a Lebesgue-integrable real function.*

The proof depends on a device used by Dunford and Morse [13, p. 419]. Let $V(t) = \sup_\Pi \sum_{i=1}^n || f(\cdot, a_i) - f(\cdot, a_{i-1}) ||$, where Π ranges over all partitions $c = a_0 \leqq a_1 \leqq \cdots \leqq a_n = t$ of the interval (c, t). The function $V(t)$ is the total variation of f on the interval (c, t), and is non-decreasing. Let $V_1(t) = V(t) + t - c$. Then V_1 is properly increasing, and $|| f(\cdot, t_1) - f(\cdot, t_2) || \leqq | V_1(t_1) - V_1(t_2) |$. The function V_1 has a derivative except on a set A of zero measure, so that if t is not in A, there is a constant M_t such that $| V_1(t + a) - V_1(t) | \leqq M_t | a |$ for every a. Consequently f has a "derivative" $\phi(\cdot, t)$ except on A; and since $|| \phi(\cdot, t) || \leqq V_1'(t)$, the function ϕ satisfies the condition of summability in Theorem 4:2.

If g lies in the set $\phi(\cdot, t)$, there exists a directed system (a_α) of real numbers such that $g = \text{w-lim}_\alpha [f(\cdot, t + a_\alpha) - f(\cdot, t)]/a_\alpha$, and such that $\lim_\alpha a_\alpha = 0$. Now let $f(x, t)$ denote the value of the functional $f(\cdot, t)$ at the point x of the space E. Then $f(x, t)$ as a function of t is of bounded variation for each x, and so has a derivative $f'(x, t)$ except on a set A_x of zero measure. It follows that if t is not in the set $A + A_x$, $g(x) = f'(x, t)$, so that $\phi(x, t) = f'(x, t)$ except on the set $A + A_x$ of zero measure. Hence $\phi(x, t)$ is integrable over (c, d) on E, since, as is well-known, the function $f'(x, t)$ is Lebesgue-integrable.

The function $V_1'(t)$ is integrable and nowhere zero, so that $\psi(\cdot, t) = \phi(\cdot, t)/V_1'(t)$ is bounded. This completes the proof of the theorem.

DEFINITION 5:3. *Let $f(t)$ be a single-valued function on the interval (c, d) to the Banach space E. Then f is absolutely continuous if for every $\epsilon > 0$ there exists an $\eta > 0$ such that $\sum_k || f(a_k) - f(b_k) || < \epsilon$ if the intervals (a_k, b_k) are disjoint and have the sum of their lengths less than η.*

It is readily verified that if the function $f(\cdot, t)$ on the interval (c, d) to E^* is absolutely continuous, it is of bounded absolute variation, and that the functions $f(x, t)$ are absolutely continuous for every x in E. It was shown in the proof of Theorem 5:1 that if $\phi(x, t)$ is the "derivative" of f, $\phi(x, t) = f'(x, t)$ almost everywhere. But if $f(x, t)$ is absolutely continuous, $f(x, t) - f(x, c) = \int_c^t f'(x, t)\, dt$. This suffices to prove the following theorem.

THEOREM 5:2. *If the function $f(\cdot, t)$ on the interval (c, d) to E^* is absolutely continuous, it has a "derivative" almost everywhere which is integrable over (c, d) on E. For each x, $\phi(x, t) = f'(x, t)$ almost everywhere, and*

$$f(\cdot, t) - f(\cdot, c) = \int_c^t \phi(\cdot, t)\, dt.$$

We now turn to the question of when the "derivative" is single-valued almost everywhere, and when it possesses stronger integrability properties.

DEFINITION 5:4. (1) *If the subset G of the adjoint E^* of the Banach space E is total, the function $f(t)$ on the interval (c, d) to E is said to have a weak derivative $f'(t)$ on G at t if $f'(t) = \text{w-lim}_{a=0} [f(t + a) - f(t)]/a$ on G.*

(2) *The function $f(t)$ on the interval (c, d) to the space E has a strong derivative $f'(t)$ at t if $f'(t) = \lim_{a=0} [f(t + a) - f(t)]/a$ in the norm.*

A preliminary observation can be made. Suppose that the "derivative" $\phi(\cdot, t)$ of a function $f(\cdot, t)$ with values in an adjoint space E^* is single-valued at a point t. Then $\phi(\cdot, t)$ is the weak derivative of $f(\cdot, t)$ on E at t. For, if $\phi(\cdot, t)$ exists, the set of numbers $[f(x, t + a) - f(x, t)]/a$ is bounded for each element x of E, so that if $\lim_n a_n = 0$, there is a subsequence (a_{n_k}) such that $\lim_k [f(x, t + a_{n_k}) - f(x, t)]/a_{n_k}$ exists. The sets

$$h\,[\text{all } [f(\cdot, t + a_{n_k}) - f(\cdot, t)]/a_{n_k} \text{ with } k \geqq m]$$

have the point $\phi(\cdot, t)$ in common, so that there is a directed set (a_α) of the numbers a_{n_k} such that $\lim_\alpha a_\alpha = 0$, and such that $\phi(\cdot, t) = \text{w-lim}_\alpha [f(\cdot, t + a) - f(\cdot, t)]/a$ on E. Consequently $\phi(x, t) = \lim_\alpha [f(x, t + a_\alpha) - f(x, t)]/a_\alpha = \lim_k [f(x, t + a_{n_k}) - f(x, t)]/a_{n_k}$, so that $\phi(x, t) = \lim_{a=0} [f(x, t + a) - f(x, t)]/a$ for every element x of E. This result is used in the proof of the next theorem, which was obtained by Gelfand [15, p. 245].

THEOREM 5:3. *If the function f on the interval (c, d) to E^* is of bounded absolute variation, and if E is separable, the "derivative" $\phi(\cdot, t)$ is almost everywhere the weak derivative of f on E.*

If furthermore the space E^* is separable, the "derivative" of f is almost everywhere the strong derivative of f. This function ϕ is then Bochner integrable, and if f is absolutely continuous, $f(\cdot, t) - f(\cdot, c) = \int_c^t \phi(\cdot, t)\, dt$ in Bochner's sense.

In view of what has just been shown, it will suffice for the first part to prove that $\phi(\cdot, t)$ is single-valued almost everywhere. If then E is separable, let (x_n) be a sequence everywhere dense on E. The function $\phi(x_n, t)$ is single-valued except on a set A_n of zero measure, by Theorem 5:2; hence $\phi(\cdot, t)$ is single-valued except on $\sum_n A_n$.

To prove the second part of the theorem, we first observe that if E is separable, $\| \phi(\cdot, t) \|$ is integrable. For, if (x_n) is the sequence everywhere dense on E, $\| \phi(\cdot, t) \| = \sup_n | \phi(x_n, t) | / \| x_n \|$ almost everywhere, so that $\| \phi(\cdot, t) \|$ is measurable. But, by Theorem 5:1, $\| \phi(\cdot, t) \|$ is bounded above by an integrable function, so that it is itself integrable. Now let $n\phi_n(\cdot, t) = \phi(\cdot, t)$

if $\| \phi(\cdot, t) \| \leqq n$, and let $\phi_n(\cdot, t) = 0$ otherwise. Then $\phi(\cdot, t) = \lim_n n\phi_n(\cdot, t)$ in the norm almost everywhere, and $\phi_n(x, t)$ is measurable for every element x of E. We denote by $\phi_n^{-1}(U)$ the set of numbers t such that $\phi_n(\cdot, t)$ lies in the subset U of E^*. Then if U is a member of the neighborhood system (1:1) used in Theorem 1:3, $\phi_n^{-1}(U)$ is measurable. But if E is separable, the sphere K of that theorem satisfies the second countability axiom, so that if U is a Borel set in the topology of Theorem 1:3, $\phi_n^{-1}(U)$ is measurable. We now note that a closed subsphere K_0 of K is weakly closed, so that K_0 is a Borel set in the weak topology; and since open subspheres of K are sums of a denumerable number of closed subspheres, these too are Borel sets in the weak topology.

If E^* is separable, every subset of K which is Borel measurable in the metric topology is a Borel set in the weak topology. Hence K can be covered by a sequence of disjoint subsets (B_k) each Borel measurable in the weak topology, and each of diameter not exceeding $1/m$. Let g_k lie in B_k, and let $\phi_{nm}(t) = g_k$ if t lies in $\phi_n^{-1}(B_k)$. Then the functions ϕ_{nm} are Bochner measurable, and $\lim_m \phi_{nm}(t) = \phi_n(\cdot, t)$ in the norm, so that the ϕ_n, and hence ϕ are Bochner measurable [10, p. 263]. Since $\| \phi(\cdot, t) \|$ is integrable, $\phi(\cdot, t)$ is Bochner integrable [10, p. 265]. But Birkhoff has shown that a Bochner integrable function is integrable in his sense to the same value [11, p. 377], and it has been indicated in Section 4 that Birkhoff integrable functions are integrable over E to the same value. Consequently, $f(\cdot, t) - f(\cdot, c) = \int_c^t \phi(\cdot, t)\, dt$ in Bochner's sense, if f is absolutely continuous. The proof is completed on noting that indefinite Bochner integrals have their integrands as their strong derivatives almost everywhere [10, p. 269].

If f is not absolutely continuous, one can apply a criterion of Dunford and Morse [13, p. 417] to show that $g = f - \int_c^t \phi\, dt$ has a strong derivative vanishing almost everywhere. If $\epsilon > 0$, there a partition $\Pi = (a_0 = c \leqq a_1 \leqq \cdots \leqq a_n = d)$ such that $V(g) \leqq \sum_i \| g(\cdot, a_i) - g(\cdot, a_{i-1}) \| + \epsilon/2^{j+1}$. But there exist x_i of unit norm such that $\| g(\cdot, a_i) - g(\cdot, a_{i-1}) \| \leqq | g(x_i, a_i) - g(x_i, a_{i-1}) | + \epsilon/n2^{j+1} \leqq V[g(x_i, \cdot), E_i] + \epsilon/n2^{j+1}$, where E_i denotes the interval (a_{i-1}, a_i). The criterion cited can be applied to the real singular functions $g(x, t)$ to yield open sets G_i with measures each not exceeding $\epsilon/n2^j$, such that $V[g(x_i, \cdot), E_i] = V[g(x_i, \cdot), G_i]$. Also, if x is of unit norm, $V[g(x, \cdot), G] \leqq V[g, G]$. Hence, if $G^j = \sum_1^n G_i$, $mG^j < \epsilon/2^j$ and $V(g) \leqq V[g, G^j] + \epsilon/2^j$; and if $G = \sum_1^\infty G^j$, $mG < \epsilon$, and $V(g) = V(g, G)$. Thus the criterion cited is satisfied, and g has a strong derivative equal to zero almost everywhere, so that f has a strong derivative as stated in the theorem.

Before proceeding further with the question of strong derivatives, we prove some lemmas on uniformly convex spaces. A Banach space E is said to be *uniformly convex* if for every ϵ between zero and two there is a $\delta(\epsilon) > 0$ such that $\| x + y \| \leqq 2(1 - \delta(\epsilon))$ whenever $\| x \| = \| y \| = 1$, and $\| x - y \| \geqq \epsilon$ [12, p. 396]. We shall prove that a uniformly convex space E is *reflexive*; that

is, if ϕ is in the second adjoint E^{**} of E, there exists an x in E such that $\phi(f) = f(x)$ for every element f of E^*.

LEMMA 5:2. *If ϕ belongs to the second adjoint E^{**} of the Banach space E, there exists a directed system (x_α) of elements of E such that $\| x_\alpha \| = \| \phi \|$, and such that $\phi(f) = \lim_\alpha f(x_\alpha)$ for every element f of E. The directed set A can be chosen to be independent of ϕ.*

This lemma is due to Goldstine [17, p. 128].

LEMMA 5:3. *If E is uniformly convex, so is E^{**}.*

Let ϕ_1, ϕ_2 be in E^{**}, and let $\| \phi_1 \| = \| \phi_2 \| = 1$. By Lemma 5:2, there exist two directed systems (x_α), (y_α), such that $\| x_\alpha \| = \| y_\alpha \| = 1$, $\lim_\alpha f(x_\alpha) = \phi_1(f)$, and $\lim_\alpha f(y_\alpha) = \phi_2(f)$. Then, if $\epsilon = \| \phi_1 - \phi_2 \|$, it is true that $\epsilon \leqq \liminf_\alpha \| x_\alpha - y_\alpha \|$. Hence there is an α_ϵ such that $\| x_\alpha - y_\alpha \| \geqq \epsilon/2$ if $\alpha R \alpha_\epsilon$ holds. But then, since $\phi_1(f) + \phi_2(f) = \lim_\alpha f(x_\alpha + y_\alpha)$, it follows that $\| \phi_1 + \phi_2 \| \leqq \liminf_\alpha \| x_\alpha + y_\alpha \| \leqq 2(1 - \delta(\epsilon/2))$, so that E^{**} is uniformly convex.

LEMMA 5:4. *Let the subset G of the space E^* be linear and total, and such that $\| x \| = \sup_{g \text{ in } G} | g(x) | / \| g \|$. Then, if E is uniformly convex, if $x = \text{w-lim}_\alpha x_\alpha$ on G, and if $\lim_\alpha \| x_\alpha \| = \| x \|$, it is true that $\lim_\alpha \| x - x_\alpha \| = 0$.*

For if not, there is an $\epsilon > 0$ and for each α an α_α such that $\| x - x_{\alpha_\alpha} \| \geqq \epsilon$, and such that $\alpha_\alpha R \alpha$ holds. It is then readily verified that $x = \text{w-lim}_\alpha x_{\alpha_\alpha}$ on G, so that it can be assumed that for all α $\| x - x_\alpha \| \geqq \epsilon$. It will also be sufficient to assume that $\| x \| = \| x_\alpha \| = 1$. By the same method as that used to prove Theorem 1:1, it can be shown that if $x = \text{w-lim}_\alpha x_\alpha$ on G, $\| x \| \leqq \liminf_\alpha \| x_\alpha \|$. Then, since $2x = \text{w-lim}_\alpha (x + x_\alpha)$ on G, it follows that $2 = \| 2x \| \leqq \liminf_\alpha \| x + x_\alpha \| \leqq 2(1 - \delta(\epsilon)) < 2$, a contradiction.

THEOREM 5:4.[2] *If the Banach space E or its adjoint E^* is uniformly convex, then E is reflexive.*

Suppose first that E is uniformly convex, and let ϕ be in E^{**}. Then, by Lemma 5:2, there exists a system (x_α) of elements of E such that $\| x_\alpha \| = \| \phi \|$, and such that $\phi(f) = \lim_\alpha f(x_\alpha)$. Let us set $\phi_x(f) = f(x)$, $\phi(f) = F_f(\phi)$, and let G be the set of all F_f. Then, by Lemma 5:3, E^{**}, G, ϕ, and the directed system (ϕ_{x_α}) satisfy the conditions of Lemma 5:4, so that $\lim_\alpha \| \phi_{x_\alpha} - \phi \| = 0$. But, if this be so, we can choose a sequence (x_n) of the elements x_α such that $\lim_n \| \phi_{x_n} - \phi \| = 0$. Consequently, the sequence (x_n) converges to an element x of E, and $\phi = \phi_x$, so that E is reflexive.

If E^* is uniformly convex, then E^* is reflexive. But Pettis [18, p. 421] has shown that this implies that E is reflexive. It should be noted that E will be reflexive if, instead of E or E^* being uniformly convex, one of E or E^* is equivalent to a uniformly convex space.

We can now summarize the results of Clarkson [12, p. 407], Pettis [18, p. 427], and Dunford and Morse [13, p. 415] in the following theorem, which is a corollary of Theorem 5:3.

[2] Proved by D. Milman, *On some criteria for the regularity of spaces of type* (B), Comptes Rendus de l' Academie des Sciences de l' URSS, vol. 20 (1938), pp. 243–246. See also B. J. Pettis, Duke Journal of Mathematics, vol. 5 (1939), pp. 249–253.

THEOREM 5:5. *Each of the following conditions is sufficient that a function f on an interval (c, d) to a Banach space E which is of bounded absolute variation should have a strong derivative almost everywhere:*

(1) *Every linear, closed, separable subspace of E is reflexive;*

(2) *E is reflexive* [18, p. 427];

(3) *E is uniformly convex* [12, p. 407];

(4) *E^* is uniformly convex;*

(5) *E has a basis (ξ_n) such that the series $\sum_n a_n \xi_n$ converges in the norm whenever its partial sums are bounded* [13, p. 415].

By Theorem 5:4, if either of E or E^* is uniformly convex, then E is reflexive. If E is reflexive, then E satisfies condition (1) [18, p. 423]. Hence to prove the sufficiency of conditions (2), (3), (4), we need only prove that (1) is sufficient.

If f is of bounded absolute variation, it is readily verified that f can have at most a denumerable number of discontinuities. It follows that the values of $f(t)$ lie in a separable subspace E_0 of E, which can be assumed to be linear and closed. If (1) holds, E_0 is reflexive, and hence one can regard f as having values in E_0^{**}. But E_0^{**} is a separable space adjoint to the space E_0^*, so that, by Theorem 5:3, f has a strong derivative almost everywhere. To prove that (5) suffices, we use Theorem 2:2. E is then separable, and is isomorphic to the adjoint E_0^* of a Banach space E_0. From Theorem 5:3 it follows that f has a strong derivative almost everywhere.

6. Representation of linear operations

In this section two representation theorems are proved. The first of these gives the form of the linear and continuous functionals on the Banach space $\mathfrak{X}$ of all Bochner integrable functions g on an interval (c, d) to a Banach space E. The norm in $\mathfrak{X}$ is defined by the equation $\| g \| = \int_c^d \| g(t) \| \, dt$ [10, p. 272]. The result given reduces to the well-known representation of linear continuous functionals on the space (L) of Lebesgue integrable functions.

THEOREM 6:1. *If the functional ξ is in $\mathfrak{X}^*$, there exists a function $\phi(\cdot, t)$ on the interval (c, d) to E^*, which is bounded and integrable over (c, d) on E, and is such that $\phi[g(t), t]$ is integrable for every g of $\mathfrak{X}$, and*

$$\xi(g) = \int_c^d \phi[g(t), t] \, dt,$$

with

$$\| \xi \| = \operatorname{ess\,sup}_t \| \phi(\cdot, t) \|$$

Let $K(t, x, r) = x$ if $t \leqq r$, and let $K(t, x, r) = 0$ otherwise. Then for each r in (c, d) and x in E, $K(\cdot, x, r)$ is in $\mathfrak{X}$ and $\| K(\cdot, x, r) \| = \| x \| \cdot | r - c |$. The function $p(\cdot, r)$ defined by the equation $p(x, r) = \xi[K(\cdot, x, r)]$ has values in E^*, and $\| p(\cdot, r) - p(\cdot, s) \| \leqq \| \xi \| \cdot | r - s |$. Hence $p(\cdot, t)$ is absolutely continuous, so that by Theorem 5:2 it has a "derivative" $\phi(\cdot, t)$ almost everywhere,

such that $\| \phi(\cdot, t) \| \leqq \| \xi \|$ almost everywhere. For each r and x, $\phi[K(t, x, r), t]$ is integrable with respect to t, and $\xi[K(\cdot, x, r)] = \int_c^d \phi[K(t, x, r), t]\, dt$. It follows that if σ is a step-function in $\mathfrak{X}$, $\xi(\sigma) = \int_c^d \phi[\sigma(t), t]\, dt$. Also $| \phi[g(t), t] | \leqq \| g(t) \| \cdot \| \xi \|$ almost everywhere. If g is in $\mathfrak{X}$, there exists a sequence (σ_n) of step-functions of $\mathfrak{X}$, such that $g(t) = \lim_n \sigma_n(t)$ in the norm of E except on a set A_0 of zero measure, and such that $g = \lim_n \sigma_n$ in the norm of $\mathfrak{X}$ [10, p. 272]. Let (x_m) be the sequence of values of the step-functions σ_n ; then $\phi(x_m, t)$ is single-valued and measurable except on a set A_m of zero measure, so that $\lim_n \phi[\sigma_n(t), t] = \phi[g(t), t]$ except on $\sum_{i=0}^{\infty} A_i$. But $|\phi[g(t), t]| \leqq \| \xi \| \cdot \| g(t) \|$ almost everywhere, so that $\phi[g(t), t]$ is integrable. Consequently $\xi(g) = \int_c^d \phi[g(t), t]\, dt$, since

$$\limsup_n \int_c^d | \phi[g(t) - \sigma_n(t), t] |\, dt \leqq \limsup_n \| \xi \| \cdot \| g - \sigma_n \| = 0.$$

To determine the norm of ξ, we note that $| \xi(g) | \leqq \operatorname{ess\,sup}_t \| \phi(\cdot, t) \| \cdot \| g \|$ and that $\| \phi(\cdot, t) \| \geqq \| \xi \|$ almost everywhere.

THEOREM 6:2. *If T is a linear and continuous transformation on the space (L) of Lebesgue integrable functions on the interval (c, d) to the adjoint E^* of a Banach space E, there exists a bounded function $\phi(\cdot, t)$ on (c, d) to E^* integrable over (c, d) on E such that for every function g in (L)*

$$Tg = \int_c^d g(t)\phi(\cdot, t)\, dt,$$

with

$$\| T \| = \operatorname{ess\,sup}_t \| \phi(\cdot, t) \|.$$

Let $K(t, r) = 1$ if $t \leqq r$, and let $K(t, r) = 0$ otherwise. Then $K(\cdot, r)$ is in (L) for every r in (c, d), and $\| K(\cdot, r) - K(\cdot, s) \| = | r - s |$, so that if $p(\cdot, r) = TK(\cdot, r)$, $\| p(\cdot, r) - p(\cdot, s) \| \leqq \| T \| \cdot | r - s |$. Hence p is absolutely continuous, and by Theorem 5:2, it has a "derivative" $\phi(\cdot, t)$ on (c, d) to E^* integrable on E, and such that $\| \phi(\cdot, t) \| \leqq \| T \|$. If g is an element of (L), $\phi(\cdot, t)g(t)$ is weakly summable, and if x lies in E, $\phi(x, t)g(t)$ is integrable. Hence $\phi(\cdot, t)g(t)$ is integrable over (c, d) on E, and if g is a step-function, $Tg = \int_c^d g(t)\, \phi(\cdot, t)\, dt$. But step-functions are everywhere dense in (L), so that if $f = Tg$, $f(x) = \int_c^d \phi(x, t)g(t)\, dt$. Hence $Tg = \int^d g(t)\phi(\cdot, t)\, dt$, and $| f(x) | \leqq \| g \| \cdot \| x \| \operatorname{ess\,sup}_t \| \phi(\cdot, t) \|$, which determines the norm of T.

COROLLARY 6:1. *If T is a linear continuous transformation on (L) to the Banach space E, there exists a bounded function $\phi(\cdot, t)$ on (c, d) to E^{**} integrable over (c, d) on E^* such that, if $x = Tg$, and $\phi_x(f) = f(x)$, then $\phi_x = \int_c^d g(t)\phi(\cdot, t)\, dt$, and $\| T \| = \operatorname{ess\,sup}_t \| \phi(\cdot, t) \|$.*

To prove this corollary, we make use of the transformation $\phi_x(f) = f(x)$, so that T can be regarded as having values in E^{**}.

The results of Theorems 5:3 and 5:5 can be applied if the range E of T satisfies conditions which are sufficient that functions of bounded absolute variation should have strong derivatives. For such transformations the function $\phi(\cdot, t)$ always has its values in E, and the integral of the representation is a Bochner integral.

In conclusion, it should be noted that the results of Section 5 can be obtained without trouble for functions of intervals in a Euclidean space; the proofs depend on the validity of these results for numerically-valued functions.

THE UNIVERSITY OF CHICAGO.

BIBLIOGRAPHY

1. Fréchet, M., *Sur l'intégrale d'une fonctionelle étendue à un ensemble abstrait*, Bulletin de la Société Mathématique de France, vol. 43 (1915), pp. 248–65.
2. Moore, E. H., and Smith, H. L., *A general theory of limits*, American Journal of Mathematics, vol. 44 (1922), pp. 102–21.
3. Hildebrandt, T. H., *On uniform boundedness of sets of functional operations*, Bulletin of the American Mathematical Society, vol. 29 (1923), pp. 309–15.
4. Alexandroff, P., and Urysohn, P., *Zur Theorie der topologischen Räume*, Mathematische Annalen, vol. 92 (1924), pp. 258–66.
5. Alexandroff, P., *Über die Struktur der bikompakten topologischen Räume*, Mathematische Annalen, vol. 92 (1924), pp. 267–74.
6. Schauder, J., *Zur Theorie stetiger Abbildungen in Funktionalräumen*, Mathematische Zeitschrift, vol. 26 (1927), pp. 47–65.
7. Orlicz, W., *Beiträge zur Theorie der Orthogonalentwicklungen*, Studia Mathematica, vol. 1 (1929), pp. 241–255.
8. Tychonoff, A., *Über die Erweiterung von Räumen*, Mathematische Annalen, vol. 102 (1930), pp. 544–61.
9. Banach, S., *Théorie des Opérations Linéaires*, 1932.
10. Bochner, S., *Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind*, Fundamenta Mathematicae, vol. 20 (1933), pp. 262–76.
11. Birkhoff, Garrett, *Integration of functions with values in a Banach space*, Transactions of the American Mathematical Society, vol. 38 (1935), pp. 357–78.
12. Clarkson, J. A., *Uniformly convex spaces*, Transactions of the American Mathematical Society, vol. 40 (1936), pp. 396–414.
13. Dunford, Nelson, and Morse, A. P., *Remarks on the preceding paper of James A. Clarkson*, Transactions of the American Mathematical Society, vol. 40 (1936), pp. 415–20.
14. Stone, M. H., *Applications of the theory of Boolean rings to general topology*, Transactions of the American Mathematical Society, vol. 41 (1937), pp. 375–481.
15. Gelfand, I., *Zur Theorie abstrakter Funktionen*, Comptes Rendus de l'Academie des Sciences de l'URSS, vol. 17 (1937), pp. 243–5.
16. Gelfand, I., *Operatoren und abstrakte Funktionen*, Comptes Rendus de l'Academie des Sciences de l'URSS, vol. 17 (1937), pp. 245–8.
17. Goldstine, H. H., *Weakly complete Banach spaces*, Duke Mathematical Journal, vol. 4 (1938), pp. 125–31.
18. Pettis, B. J., *A note on regular Banach spaces*, Bulletin of the American Mathematical Society, vol. 44 (1938), pp. 420–8.